Ricardian selection

Finicelli, Andrea, Pagano, Patrizio and Sbracia, Massimo
Bank of Italy

August 2009

Ricardian Selection*

Andrea Finicelli, Patrizio Pagano, and Massimo Sbracia

Bank of Italy, Via Nazionale 91, 00184 Rome, Italy

Abstract

We analyze the foundations of the relationship between trade and TFP in the Ricardian model. Under general assumptions about the autarky distributions of industry productivities, trade openness raises TFP. This is due to the selection effect of international competition — driven by comparative advantages — which makes "some" high- and "many" low-productivity industries exit the market. We derive a model-based measure of this effect that requires only production and trade data. For a sample of 41 countries, we find that Ricardian selection raised manufacturing TFP by 11% above the autarky level in 2005 (6% in 1985), with a neat positive time trend and large cross-country differences.

JEL classification: F10, D24, O40

Keywords: selection effect, Eaton-Kortum model, international competition

* This paper supersedes the first part of a paper that we previously circulated as "The selection effect of international competition." We thank Jonathan Eaton, Dirk Krueger, Mark Roberts, Andrés Rodríguez-Clare, Fabiano Schivardi, Bas Straathof, Lucia Tajoli, and Daniele Terlizzese for helpful suggestions, and seminar participants at the Bank of Italy, Einaudi Institute for Economics and Finance (EIEF), European Commission, NBER PRBB Summer Institute 2007, SED Meeting 2008, University Luiss Guido Carli, and University of Cagliari (CRENoS Conference) for comments. We are indebted to several colleagues, and especially Paola Caselli, Alberto Felettigh, Marcello Pagnini, Massimiliano Pisani, and Enrico Sette, for many constructive discussions. Giovanna Poggi provided valuable research assistance. All the remaining errors are ours. The paper was completed while Sbracia was visiting EIEF, whose fruitful hospitality is gratefully acknowledged. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Italy. E-mail: andrea.finicelli@bancaditalia.it, patrizio.pagano@bancaditalia.it, massimo.sbracia@bancaditalia.it.

1 Introduction

Few economic theories have been studied as extensively as the Ricardian model of international trade, which is now almost two centuries old. Nonetheless, within this model the relationship between trade openness and Total Factor Productivity (TFP) has generally been overlooked. One possible reason is that the standard model without externalities and distortions, while implying that trade is always welfare improving, delivers ambiguous predictions about the effect of trade on TFP. In particular, it is possible to construct simple examples in which one country holds a comparative advantage in the production of low-productivity goods, so that its TFP declines after removing trade barriers. Yet, growing and robust empirical evidence — especially studies based on firm-level data — points out that trade has a significant positive impact on TFP.[1] This raises some intriguing questions: is there any key feature of the open economy missing in the Ricardian model? Or, rather, are the examples in which TFP does not rise so special that they can be safely disregarded?

In this paper we provide a thorough analysis the relationship between trade and TFP in the Ricardian model and show that the latter interpretation is correct. We address this issue building on the most general version of the model, with many countries and a continuum of goods, developed by Eaton and Kortum (2002; EK hereafter). We show that trade openness may reduce TFP *only if* industry productivities (country technologies) in different countries are highly correlated *and* their joint distribution does not belong to the families that are most commonly used in the literature, such as the Fréchet, Pareto, normal, and lognormal. In contrast, if industry productivities in different countries are independent, trade openness raises TFP *for any distribution of technologies*; for the families of distributions mentioned above, TFP rises *for any degree of correlation.*

Key to these results is the fact that the selection effect of international competition favors the survival of industries with, on average, higher productivity. We show that the comparison between the TFP under autarky and the TFP of an open economy boils down to a comparison between a simple mean and a conditional mean, where the conditioning event — that domestic industries survive international competition — tends to lift TFP after opening to trade.

Examples in which international competition induces an "adverse" selection in favor of industries with low productivity can still be constructed, but they crucially require a high positive correlation among country technologies. The continuum-of-good assumption and the probabilistic representation of productivities unveil that comparative advantages are

[1] Among the most influential papers see Bernard and Jensen (1999), Frankel and Romer (1999), Pavcnik (2002), Bernard, Eaton, Jensen, and Kortum (2003), Dollar and Kraay (2003), and Alacalá and Ciccone (2004). For a recent survey with an emphasis on firm-level data see Bernard, Jensen, Redding, and Schott (2007).

inversely related to the correlations between technologies. High correlations correspond, for any country, to low comparative advantages, a small TFP gain from trade (i.e. the ratio between the open economy's and the autarky's TFPs), and a narrow scope for international trade. The result that under independence TFP always raises is just a consequence of this insight. For the families of distributions mentioned above correlation still matters, because it determines the size of the TFP gain from trade, but the latter is always positive and vanishes as correlation becomes perfect.

With Fréchet distributed technologies as in EK, we also characterize the TFP of the tradeable sector in an open economy with a very compact expression that sheds light on its determinants. An increase in TFP may be due to "genuine" domestic technological progress, or it may reflect other factors such as an improvement in the technologies of competitor countries, loosening trade barriers (including the entry of new competitors), declining foreign input costs, or rising domestic input costs. All these factors raise the TFP through the selection effect, by affecting the share of low-productivity industries that are forced to exit. TFP gains from trade are also increasing in the dispersion of industry productivities that, together with correlation, is related to the extent of comparative advantages.

Throughout the paper, we define the TFPs of closed and open economies as the first moments of the corresponding distributions of industry productivities. If consumers have CES preferences and productivities follow the most common distributions used in the literature, we show that this definition is appropriate in the extreme cases of autarky and absent trade barriers. The average of industry productivities weighted by the value of industry outputs, in fact, is proportional to the first moment of the distribution. The proportionality constant gathers the whole effect of consumer preferences — with more substitutability among goods increasing the weight of high-productivity industries and raising the weighted TFP — and cancels out when we compute the TFP gain of free trade *versus* autarky. This gain reflects a "pure" selection effect, as the weights of the industries that eventually die are reallocated proportionally to the surviving ones. In the intermediate case of positive but finite trade barriers, the TFP gain from trade features an additional effect, that has always a non-negative sign, given by the reallocation of market shares towards exporters.

Focusing on the selection effect, we also provide a result that has compelling empirical implications. We show that the TFP of the open economy is equal to the autarky TFP, augmented by a measure of trade openness that requires only data on production and trade. Thus, the selection effect can be easily quantified. We perform such measurement for a sample of 41 countries with annual data in the period 1985-2005. Averaging across countries, we find that in 1985 international competition raised manufacturing TFP by 6 percent above the autarky level.[2] Not surprisingly, due to enhanced trade integration, this contribution

[2]When we bring the model to the data, our definition of tradeable sector boils down to the manufacturing sector.

exhibits a neat positive time trend, shared by most countries, growing to as much as 11 percent in 2005. In the cross-section, however, TFP gains from trade vary widely.[3]

Our paper is closely related to Melitz (2003) (and the subsequent literature, including Chaney, 2008, and Melitz and Ottaviano, 2008), who also derives a positive relationship between trade and TFP.[4] There are, however, two key differences. First, we obtain the result *with perfect competition*, whereas Melitz assumes monopolistic competition. Second, here *selection is driven by comparative advantages*, as is to be expected within the Ricardian approach. Therefore, not only low- but also high-productivity domestic industries can exit the market (and be replaced by imports from foreign industries), although with lower probabilities, so that removing trade barriers generates some "action" along the whole distribution of productivities. In Melitz, instead, all and only the firms whose productivity is below a certain threshold exit the market after trade barriers decline.[5] To stress these differences, we describe this mechanism with the expression *Ricardian selection* instead of *self-selection*. The latter denomination is common in the monopolistic competition literature, where low-productivity firms really self-select by refraining from producing or exporting whenever they expect negative profits. In the Ricardian model, instead, international competition forces both low- and high-productivity industries to exit.

Our findings also bring this paper close to the literature that emphasizes the role of institutions (or "social infrastructure", as in Hall and Jones, 1999) in explaining TFP differences across countries. Examples include Conway and Nicoletti (2006) and Lagos (2006), who show that higher regulation in the non-tradeable sector and in the labor market lowers the TFP of the tradeable sector. Our analysis shows, in contrast, that higher regulation, by rising domestic costs and forcing low-productivity industries to exit, has the opposite effect. In addition, the role of other factors, such as proximity to high-TFP countries (in other words, *geography*), also emerges.

The rest of the paper is organized as follows. Section 2 offers a brief outline of the EK

[3]Our analysis focuses on the effect of trade on the TFP of *the tradeable sector only* — an effect that the Ricardian model allows to express analytically and quantify. We discuss implications for the TFP of the whole economy, however, when it is pertinent.

[4]Other close relatives of this paper are Bernard, Eaton, Jensen, and Kortum (2003) and Waugh (2008). The former analyzes trade and productivity with Bertrand competition, but does not derive a closed-form expression for the aggregate TFP. The latter builds a variant of the EK model with traded intermediate goods and non-traded final goods and measures the contribution of trade to cross-country income differences. This contribution turns out to be small, but this is due to the fact that poor countries face higher costs of exporting then rich countries; removing these asymmetries would lower income differences by about a half.

[5]With the assumption of constant returns to scale, Ricardian models leave the size of firms indeterminate, as the equilibrium pins down only the size of the whole industry. Hence, the comparison between monopolistic competition and Ricardian models is loose. The former, in fact, focuses on firms that produce heterogeneous goods (and analyzes intra-industry trade), while the latter focuses on industries, with an industry defined as the firm or set of firms (with identical productivities) that produces an homogeneous good.

model. Section 3 presents our main theoretical results about trade and TFP. In Section 4 we elaborate on our results, providing some intuition and extending them to more general distributional assumptions. In Section 5 we analyze the weighted TFP in closed and open economies. In Section 6 we quantify the magnitude of the selection effect on the TFP of the manufacturing sector. Section 7 concludes.

2 An outline of the Eaton-Kortum model

EK consider a Ricardian framework with N countries ($N > 1$) and a continuum of tradeable goods produced with constant-returns-to-scale technologies. Denote by $z_i(j) > 0$ the efficiency of country i in producing the tradeable good j, with $i \in \{1, ..., N\}$ and $j \in [0, +\infty)$; namely: $q_i(j) = z_i(j) \cdot I_i(j)$, where $q_i(j)$ is the amount of good j produced by country i and $I_i(j)$ is the bundle of inputs, which combines labor and intermediate goods, needed to produce that output.

The key hypothesis is that each $z_i(j)$ is the realization of a country-specific random variable Z_i. Specifically, it is assumed that for any country i: $Z_i \sim \textit{Fréchet}(T_i, \theta)$, with $T_i > 0$, $\theta > 1$, and $\{Z_i\}_{i=1}^N$ mutually independent. Due to the continuum-of-goods assumption and assuming the law of large numbers, the share of goods for which country i's efficiency is below any positive real z is simply the probability $\Pr(Z_i < z) = F_i(z) = \exp\left(-T_i \cdot z^{-\theta}\right)$, where F_i denotes the cumulative distribution function (c.d.f.) of Z_i.[6]

EK show that the parameters T_i and θ are the theoretical counterparts, in a context with many countries and a continuum of goods, of the Ricardian concepts of absolute and comparative advantages. T_i, to which we will refer as *state of technology*, captures country i's absolute advantages: an increase in T_i relative to T_n implies an increase in the share of goods that country i produces more efficiently than country n. In turn, θ is inversely related to the dispersion of Z_i and is a measure of the *precision* of the distribution.[7] Its connection with the concept of comparative advantage stems from the fact that, in Ricardo, gains from trade depend on heterogeneities in technologies across countries. In this perspective, EK demonstrate that a decrease in θ (i.e. higher heterogeneity in each country), coupled with mutual independence, generates larger gains from trade for all countries. The extension of our results to correlated distributions, presented in Section 4, is also helpful to clarify that

[6] Kortum (1997) and Eaton and Kortum (2009) show that the Fréchet distribution emerges from a dynamic model in which, at each point in time: (*i*) the number of ideas that arrive about how to produce a good follows a Poisson distribution; (*ii*) the efficiency conveyed by each idea is a random variable with a Pareto distribution; (*iii*) firms produce goods using always the best idea that has arrived to them. Jones (2005) shows that this set up on the flow of ideas entails two other results: the global production function is Cobb-Douglas and technical change in the long run is labor-augmenting.

[7] T_i and θ are related to both the mean and the variance of Z_i. Denoting Euler's Gamma function by Γ, the moment of order k of Z_i, which exists only if $\theta > k$, is $T_i^{k/\theta} \cdot \Gamma[(\theta - k)/\theta]$.

comparative advantages are related to both *the heterogeneity of technologies within countries and the correlation of technologies between countries.*

A second set of assumptions concerns input costs and trade barriers. The cost of the bundle of inputs in country i is denoted by c_i and it is split into wages and prices of intermediate goods. Trade barriers are modeled as Samuelson's iceberg costs: delivering one unit of good from country i to country n requires producing d_{ni} units, with $d_{ni} > 1$ for $i \neq n$ and $d_{ii} = 1$ for any i. By arbitrage, trade barriers obey the triangle inequality, so that $d_{ni} \leq d_{nk} \cdot d_{ki}$ for any n, i and k.

As for the market structure, the model assumes perfect competition. Together with the hypotheses on costs and technologies, perfect competition implies that the price of one unit of good j produced by country i and delivered to country n is: $p_{ni}(j) = c_i d_{ni}/z_i(j)$. In country n each good j is purchased from the country that provides it at the lowest price, i.e.:

$$p_n(j) = \min_{i=1,\ldots,N} \{p_{ni}(j)\} \ .$$

Consumers maximize a standard CES utility function, with elasticity of substitution given by $\sigma > 1$, subject to the usual constraint that total spending cannot be larger than total income.

With this set of assumptions, EK prove two fundamental properties of the model. First, the market share of country i in country n — i.e. the ratio between the value of the imports of country n from country i (X_{ni}) and the value of the total expenditure (or total absorption) of country n (X_n) — is given by:

$$\frac{X_{ni}}{X_n} = \frac{T_i \cdot (c_i d_{ni})^{-\theta}}{\Phi_n} \ , \text{ where: } \Phi_n = \sum_{k=1}^{N} T_k \cdot (c_k d_{nk})^{-\theta} \ . \tag{1}$$

This share is increasing in the state of technology T_i and decreasing in the input cost c_i and the trade barrier d_{ni}. Its value depends also on the technologies, costs and trade barriers of any other country k: it increases with costs c_k and distances d_{nk}, and decreases if technologies T_k increase.

Second, the exact price index of the bundle of tradeable goods in country n resulting from the CES aggregator and the prices $p_n(j)$ is:

$$p_n = \gamma \cdot \Phi_n^{-1/\theta} \ , \text{ where: } \gamma = \left[\Gamma\left(\frac{\theta + 1 - \sigma}{\theta}\right)\right]^{1/(1-\sigma)} \ , \tag{2}$$

with Γ denoting Euler's Gamma function and $\theta > \sigma - 1$.

This setup is completed with two further assumptions. The first is that intermediate inputs comprise the full set of tradeable goods aggregated with the CES function with elasticity σ. Denoting by $\beta \in (0, 1)$ the constant labor share, c_i is given by:

$$c_i = w_i^{\beta} p_i^{1-\beta} \ , \tag{3}$$

where w_i is the nominal wage in country i and p_i is given by equation (2). The second hypothesis is that there is also a non-tradeable sector in the economy; thus, market shares, prices, and wages defined above are referred to the tradeable sector only.

These two further assumptions enable EK to solve the model for equilibrium prices and quantities in two polar cases. In one case, labor is mobile between the tradeable and non-tradeable sectors; in the other, it is immobile. In both cases, it is assumed that a constant fraction $\alpha \in (0,1)$ of the aggregate final expenditure is spent on tradeable goods. The solution of the model, then, is given by a system of non-linear equations, with parameters d_{ni}, T_i, θ, α and β (see EK, pp. 1756-1758).[8] Because of non-linearities, there is no closed-form solution. Nevertheless, it is possible to simulate the model and analyze some counterfactuals or rearrange the main equations in order to obtain testable implications. In the following, we keep on building on the theoretical model and show how we can use it in order to derive a theoretical expression for the TFP of tradeables.

3 TFP in open economy

In general, the TFP of a country is the average productivity across the industries that are actually engaged in production. Under autarky, production occurs over the entire range of goods. Then, the TFP of the tradeable sector of country i is the unconditional mean of Z_i, i.e. the mean across *all tradeable goods*. In the open economy, instead, any country produces only the goods in which it holds a comparative advantage. In this case, the TFP of the tradeable sector of country i must be obtained by computing the mean of Z_i only across *the goods that are actually produced by domestic industries*, thus excluding the goods that are imported because i is not sufficiently competitive.

In order to find out the productivity distribution for the industries that are sufficiently productive as to survive foreign competition, we can resort to the model. Denote the random variable that describes the productivities of the surviving industries by $Z_{i,o}$, where the subscript o stands for the open economy; its c.d.f. is:

$$F_{i,o}(z) \equiv \Pr(Z_{i,o} < z) = \Pr\left(Z_i < z | P_{ii} = \min_k P_{ik}\right) , \tag{4}$$

where P_{ik} is the random variable that describes the prices $p_{ik}(j)$ for any i and k (including $i = k$). Equation (4) establishes that the goods j eventually produced by country i are all and only those for which $p_{ii}(j) \leq p_{ik}(j)$ for any k. This point requires a formal proof. On one hand, if j is such that $p_{ii}(j) \leq p_{ik}(j)$ for any k, then country i can sell j with the lowest

[8] Alvarez and Lucas (2007) generalize the model by considering distinct final and intermediate goods, and distinguishing between tariffs and transport costs. Then, they provide sufficient conditions for existence and uniqueness of the equilibrium.

price on the domestic market and, therefore, will certainly produce it.[9] On the other hand, if $p_{ii}(j) > p_{ik}(j)$ then country i cannot charge the lowest price for j on the domestic market and, as a consequence of the triangle inequality, it is unable to provide the best deal on foreign markets either, thus it will not produce j at all (this intuitive step is formally proven in Appendix A). Computing $F_{i,o}$ yields the following result:

Proposition 1 *If technologies are Fréchet distributed and markets for tradeable goods are perfectly competitive, then:*

$$Z_{i,o} \sim \textit{Fréchet}\,(\Lambda_i, \theta) \ ,$$

where

$$\Lambda_i = T_i + \sum_{k \neq i} T_k \left(\frac{c_k d_{ik}}{c_i} \right)^{-\theta} \ . \tag{P1}$$

Proof. See Appendix A.1 ■

Thus, the variable $Z_{i,o}$ is Fréchet distributed, with mean:

$$E\left(Z_{i,o}\right) = \Lambda_i^{1/\theta} \cdot \Gamma\left(\frac{\theta - 1}{\theta}\right) \ . \tag{5}$$

$E\left(Z_{i,o}\right)$, a monotone function of Λ_i, provides a theoretical expression for the TFP of the open economy (denoted by $\text{TFP}_{i,o}$), while $E\left(Z_i\right)$ is the TFP under autarky (TFP_i). Therefore, we define the TFP gain from trade as the ratio:[10]

$$\frac{\text{TFP}_{i,o}}{\text{TFP}_i} = \frac{E\left(Z_{i,o}\right)}{E\left(Z_i\right)} = \left(1 + \sum_{k \neq i} \frac{T_k}{T_i} \left(\frac{c_k d_{ik}}{c_i} \right)^{-\theta} \right)^{1/\theta} \ .$$

Note that, although we compute the TFP of country i by averaging $z_i(j)$ across goods j, we are not summing up "apples and oranges". While our measure of TFP depends on the physical units chosen to measure the different goods, *we define these units in the same way as they enter the utility function.*[11] In particular, with CES preferences, we are averaging productivities measured in physical units that enter the utility function in a completely symmetric way.[12]

[9] Given the continuity of the random variables considered here (i.e. of Z_i and, as a consequence, of P_{ik}), we can neglect events of the type $p_{ii}(j) = p_{ik}(j)$, since they have zero probability.

[10] In a related paper, Finicelli, Pagano, and Sbracia (2009a) build on this theoretical result in order to measure TFPs of the tradeable sector of 18 OECD countries, relative to that of the United States.

[11] Summing the $z_i(j)$ across different goods is an essential operation to solve the EK model for equilibrium prices and quantities. Think for example of the price index p_i, which is nothing but an average of the $z_i(j)$ (see, for instance, equation (21)).

[12] Demidova and Rodríguez-Clare (2009) provide a very elegant treatment of this issue. In particular, they define the quantities $q_i(j)$ that enter the utility function as: $q_i(j) = \lambda_i(j)\,\tilde{q}_i(j)$, where $\tilde{q}_i(j)$ is the "raw quantity" of good j and $\lambda_i(j)$ is a function that accounts for of the different ways in which $\tilde{q}_i(j)$ can be measured. Hence, changing the unit of measurement for $\tilde{q}_i(j)$ determines a corresponding change in $\lambda_i(j)$, so that $q_i(j)$ remains unaltered.

The main implication of Proposition 1 is that $\Lambda_i > T_i$, therefore $E(Z_{i,o}) > E(Z_i)$.[13] In other words, the model predicts that the TFP of the open economy is larger that the TFP under autarky. The next section is entirely devoted to explaining and extending this result. Before that, it is worth noting that the proposition also sheds light on the factors that affect the TFP in an open economy.

Equation (P1) shows that, in an open economy, Λ_i depends not only on T_i, but also on the technologies, costs, and trade barriers of all the other countries, as well as on domestic costs. This result can be readily explained. Suppose that T_k increases for some $k \neq i$. Country k, then, produces and exports more goods than before (equation (1)), partly crowding out production in country i. The productivity of the goods that keep being produced in country i, however, is higher, on average, than the one of the goods whose production has been displaced, hence the fact that $\Lambda_i > T_i$. The effects of domestic input costs and of iceberg costs follow the same mechanics. A larger c_i crowds out production in country i in favor of its competitors, but its own average productivity increases (by the same token, the effect of an increase in foreign input costs c_k on country i's average productivity is opposite). Higher d_{ik} narrow the range of goods imported by country i, keeping alive industries with a lower average productivity. Note that as d_{ik} go to $+\infty$ for all $k \neq i$ — i.e. as the country tends to autarky — then Λ_i tends to T_i.

The positive relationship between aggregate productivity and domestic costs contrasts with the results of Lagos (2006) and Conway and Nicoletti (2006). In the EK model, if a country pays higher wages or incurs larger costs because of distorted labor or non-tradeable product markets, then the selection effect of international competition forces inefficient industries to exit, raising aggregate productivity.[14] On the contrary, in Lagos and in Conway and Nicoletti distorted markets cause an adverse selection of productive units, hampering the efficiency of their allocation and, in turn, reducing aggregate productivity. Assessing the net effect of these distortions on TFP, then, remains essentially an empirical question.[15]

[13] By the properties of the Fréchet distribution, the stronger implication that $Z_{i,o}$ first-order stochastically dominates Z_i follows immediately.

[14] Note that this improvement of TFP comes together with fewer exporters and lower market shares. It should also be clear that the larger deadweight losses associated to, e.g., higher firing costs in the labor market or lower degree of competition in the non-tradeable sector imply that increasing domestic costs *is not* a recommended policy prescription to raise TFP. More in general, these examples provide further evidence of the fact that, in an open economy, an increase (decrease) in TFP is not necessarily associated with an increase (decrease) in welfare (see also footnote 16).

[15] Chari, Restuccia, and Urrutia (2005) focus on another mechanism through which more frictions in the labor market raise the 'measured' TFP (proxied by income per worker). In their paper, the result occurs because higher firing costs increase the level of training that firms provide to workers, raising human capital and, in turn, the measured TFP. They also provide *prima facie* evidence that the relationship between the level of employment protection and the TFP across European countries is positive. Our results, then, provide an alternative explanation of the same findings based on the role of international competition.

By recalling the expressions of costs (equation (3)) and prices (equation (2)), Proposition 1 also shows that changes in technologies, costs and trade barriers do not have only a "direct" selection effect on TFP. International competition yields also second- and higher-order effects via changes in input costs. Consider, for instance, an increase in the foreign technology T_k. The increase in T_k, by making available cheaper goods in country k, lowers also its input costs c_k further enhancing its external competitiveness and providing an additional boost to the TFP of country i. This effect is partly offset by the availability of cheaper inputs in country i (i.e. by a decline in c_i), but is reinforced by lower input costs in countries other than i and k.[16]

Proposition 1 also shows that the benefits of technological progress in one country are not spread evenly on the TFP of other countries. The extent to which TFP changes following a change in foreign technologies and costs reflects the size of domestic relative to foreign costs and, inversely, that of domestic trade barriers. For instance, an increase in the technology of the United States will have a stronger (weaker) impact on closer (more distant) countries. By the same token, since the TFP in country i changes as trade barriers change, equation (P1) suggests that looking at the dynamics of TFP growth could misrepresent the picture about "genuine" technological developments during periods in which countries liberalize or place restrictions on international trade.

Equation (P1) is theoretically appealing but also rather difficult to apply in empirical studies, since it requires data on technologies, costs, and trade barriers for all countries. However, a very helpful expression for Λ_i can be derived by exploiting the fact that countries' technologies, costs, and trade barriers combine uniquely into the geographical distribution of production and trade data. In particular, we can prove that:

Proposition 2 *Given the equations for market shares (1) and costs (3), then:*

$$\Lambda_i = T_i \left(1 + \sum_{k \neq i} \frac{X_{ik}}{X_{ii}} \right) = T_i \left(1 + \frac{IMP_i}{PRO_i - EXP_i} \right) . \tag{P2}$$

Proof. See Appendix A.2 ■

Hence, Λ_i is equal to T_i augmented by a factor that depends on the ratio between the value of country i's aggregate imports (IMP_i) and the value of its production (PRO_i) net of aggregate exports (EXP_i). Let us write:

$$\Omega_i = 1 + \frac{IMP_i}{PRO_i - EXP_i} \; ; \tag{6}$$

[16] In the version considered here, the model ignores the possibility of technology spillovers across countries. In fact, the Z_i's are independent random variables and the T_i's can change in an unrelated fashion. Rodríguez-Clare (2007) extends the model to account for international diffusion of ideas. A similar route would be to consider correlated Z_i's (see the next section).

Ω_i is the ratio between country i's total absorption (or total domestic demand) and its production sold domestically. Therefore, Ω_i is a measure of *trade openness* for country i. Note that, consistently with equation (P1), as d_{ik} go to $+\infty$ for all $k \neq i$ then imports and exports go to zero and Λ_i tends to T_i.

Proposition 2 provides an interesting contribution to the literature concerning the measures of trade openness. Papers exploring the relationship between trade and productivity typically measure trade openness as the sum of nominal imports and exports scaled by the nominal GDP (*nominal openness*). An exception is Alcalá and Ciccone (2004) who scale nominal imports and exports with the GDP in PPP US dollars (*real openness*), on the ground of theoretical motivations. Our analysis finds that the Ricardian trade theory suggests to measure trade openness with Ω_i. Equation (P2), in fact, shows that Ω_i is the trade-related variable that summarizes the effects of international competition on TFP. By comparing equation (P2) with equation (P1), it is evident that Ω_i takes into account the factors related to domestic and foreign costs that are considered by Alcalá and Ciccone.

The wide availability of production and trade data makes it easy to compute Ω_i and quantify the magnitude of the selection effect for several countries and years. Before turning to the empirical analysis, however, we focus on the prediction that openness raises TFP, providing further insights about how and why this happens and exploring possible extensions of this result.

4 Intuition and extensions

The main implication of Proposition 1 is that TFP always rises when trade barriers are removed. This is a remarkable difference with respect to previous Ricardian models, where the law of comparative advantage may lead a country to specialize in the production of low-productivity goods, a situation in which aggregate TFP would diminish after opening to trade.

To build an intuition about our new result, let us retain only the essential ingredients of the model. Consider a simple case with two countries (n and i), no trade barriers (i.e. $d_{ni} = d_{in} = 1$), no intermediate goods ($\beta = 1$), and identical input costs (i.e. $c_n = c_i = 1$).[17] These assumptions simplify the distribution of the productivities of the surviving industries (that we have computed for the general case in Proposition 1), as they imply that country i produces good j if and only if $z_i(j)/z_n(j) \geq 1$. In addition, with no trade barriers producers

[17] Even though in the model costs are endogenous (which would prevent us from setting arbitrary values for c_i and c_n), we can build examples that yield $c_n = c_i = 1$. An obvious way is to assume perfectly symmetric countries where, in particular, $T_i = T_n$. However, even $T_i \neq T_n$ could still yield $c_n = c_i = 1$ if, for instance, the sizes of the labor force in the two countries are different. We strongly stress, however, that these simplifying assumptions are by no means necessary for the arguments made in this section.

and exporters coincide, making our task of providing some intuition easier.

The first important feature of the mechanism through which selection raises TFP in the Ricardian model is that any industry can survive or die after openness, and the probability that each industry survives (dies) is increasing (decreasing) in its own productivity. In fact, using both mutual independence and Fréchet distribution of technologies, the probability of surviving for an industry of country i that has a productivity z is simply:

$$\Pr\left(\frac{Z_i}{Z_n} \geq 1 | Z_i = z\right) = \exp\left(-T_n \cdot z^{-\theta}\right) \quad \text{for } z > 0 \ ,$$

which is always included in the open interval $(0, 1)$ and strictly increasing in z. (Its complement to 1, the probability that the industry dies, is always decreasing in z.) This is an interesting difference with respect to the model of Melitz (2003) in which that probability is either 0 or 1, depending on whether the firm's productivity is below or above the threshold that separates incumbents from entrants. The reason for this difference is that the EK model is governed by the law of comparative advantage. Therefore, any industry, even a high-productivity one, can exit the market and it does so if the good that it produces is made more efficiently in the rival country; this happens, however, with a probability that is lower for higher-productivity industries. Specularly, even a very low-productivity domestic industry survives if its own good is made more efficiently than in the other country — but its probability of surviving is lower than for higher-productivity industries.

Second, the selection effect makes the model consistent with the "exceptional export performance" documented by Bernard and Jensen (1999). Let us temporarily re-introduce trade barriers (otherwise all producers would also export). Good j is made in country i if and only if $z_i(j)/z_n(j) \geq d_{in}^{-1}$. In addition, if $z_i(j)/z_n(j) \geq d_{ni}$, then the good j is also exported by country i to country n (otherwise, the good j is sold only domestically). With mutually independent and Fréchet distributed technologies, and following steps similar to those illustrated in Appendix A to prove Proposition 1, we can show that the distribution of the productivities of exporters is Fréchet, with state $T_i + T_n \cdot d_{ni}^{\theta}$ and precision θ. Applying Proposition 1 to this simplified setup, we find that the distribution of the whole set of surviving industries is Fréchet with state $T_i + T_n \cdot d_{in}^{-\theta}$ and precision θ. Since both d_{ni} and d_{in} are larger than 1, then the average productivity of exporters is higher than the TFP of the whole economy (the latter being the average productivity across all the industries that survive international competition, i.e. exporters and producers that sell only domestically).[18] As in monopolistic competition models, the reason why exporters are, on average, more productive is that their goods have to be competitive enough to overcome trade barriers. However, analogously to what discussed above, the mechanisms behind selection in the two models are not identical. In Melitz, exporters and non-exporters are separated by a productivity threshold; therefore, even the worst exporter has always a higher productivity

[18] It follows trivially that the average productivity of exporters is also higher than the average productivity of non-exporters.

than the best non-exporter. In the Ricardian model, instead, as a consequence of the law of comparative advantage, few "bad" exporters and "good" non-exporters coexist with many "good" exporters and "bad" non-exporters.

Are these predictions robust to the distributional assumptions? Let us go back to the simplified framework with no trade barriers. The main result that TFP rises after removing trade barriers can be formally written as:

$$E\left(Z_i|Z_i \geq Z_n\right) \geq E\left(Z_i\right) \ . \tag{7}$$

Inequality (7) makes it clear that the comparison between the TFP of an open economy and the TFP under autarky boils down to a comparison between a conditional mean and a simple mean. The conditioning event is that domestic industries are better than foreign industries (or "sufficiently better", if there are trade barriers and heterogeneous input costs). This condition is what "tends" to raise TFP after opening to trade.

However, (7) does not hold for all possible joint distributions of Z_i and Z_n. A simple way to build a counterexample in which (7) is not satisfied is the following. Take any random variable Z_i, and construct a variable Z_n such that when Z_i takes values higher (lower) than its own mean, then Z_n is higher (lower) than Z_i. In this case, $Z_i \geq Z_n$ only for values of Z_i lower than its unconditional mean, therefore $E\left(Z_i|Z_i \geq Z_n\right) < E\left(Z_i\right)$.[19]

The key feature of this counterexample is that Z_i and Z_n are not independent, but positively correlated. This positive correlation is an obstacle that hampers TFP growth after opening to international trade. The continuum-of-goods assumption and the probabilistic representation of productivities, then, unveil the role of correlation in determining the sign and the size of TFP growth — a role that is crucial also in Ricardian models where those two characteristics are absent.

Consider for instance a two-country two-good model and let us play with the statement "country i holds a comparative advantage in the production of *low-productivity good* 1", a condition under which the TFP of country i declines. By labeling good 1 as "low productivity", we are implicitly assuming that both countries make good 1 with a lower productivity than good 2. (If 1 was a low-productivity good only for country i, this country could not have a comparative advantage in making that good.) Since $z_i\left(1\right) < E\left(Z_i\right) < z_i\left(2\right)$ and $z_n\left(1\right) < E\left(Z_n\right) < z_n\left(2\right)$, then technologies in the two countries are positively correlated.

The previous counterexample, which featured more than just two goods, was obtained simply by coupling positive correlation with the condition that one country is systematically

[19] Note that after opening to international trade, welfare increases also in countries where TFP declines. The reason is that a lower TFP is more than compensated by lower consumer prices. Demidova and Rodríguez-Clare (2009) show examples in which, by the same token, welfare falls despite higher productivity. Using a monopolistic competition model for a small economy, they also decompose welfare gains into four factors: productivity, terms of trade, product varieties, and heterogeneity across varieties.

better than the other in making low-productivity goods. With this case in mind, consider now the effect of imposing independence on the distributions of Z_i and Z_n. Independence has two main consequences: (i) it shrinks the set of goods that both countries make with low productivity; (ii) it prevents one country from being systematically better than the other in producing these goods.[20]

Thus, not surprisingly, independence between Z_i and Z_n is a sufficient condition for (7) to hold, irrespectively of the shape of the distribution of Z_i and Z_n (see Appendix C, for details). In other words, under mutual independence TFP always rises after opening to trade. In particular, the result holds for all the distributions, like Pareto, normal, lognormal, and, of course, the Fréchet, that are commonly used to describe productivities at the industry or firm level and that entail very simple analytic solutions for this model.[21]

While independence is sufficient for TFP to increase, it is by no means necessary. More importantly, we show that TFP rises irrespectively of the degree of correlation between country technologies for some multivariate families that yield, as marginals, the aforementioned Fréchet, Pareto, normal, and lognormal. The intuition is that assuming such families of distributions prevents one country from being systematically better than the other in producing low-productivity goods, even though the set of these goods can be quite "large" (i.e. the correlation can be close to 1). It holds, however, that the TFP gain is lower for higher correlations.[22]

Let us consider the multivariate Fréchet case in detail, with a simple extension that covers all levels of dependence, from independence to perfect correlation. Suppose that the random vector (Z_i, Z_n) has the following c.d.f.:

$$\Psi_{i,n}(z_i, z_n) = \exp\left\{-\left[\left(T_i \cdot z_i^{-\theta}\right)^{1/r} + \left(T_n \cdot z_n^{-\theta}\right)^{1/r}\right]^r\right\} , \tag{8}$$

where $\Psi_{i,n}(z_i, z_n) = \Pr(Z_i < z_i, Z_n < z_n)$ and $r \in (0, 1]$. This distribution yields two Fréchet as marginals (with parameters respectively equal to (T_i, θ) and (T_n, θ)) and is suggested in EK for an extension of their model to correlated technologies.[23] The parameter r is an "index

[20]By the same token, negative correlation would imply that when a country makes a good with a low productivity, then the other makes it with a high productivity, leading to even larger TFP gains in both countries. In the following, we show this result formally in the case of normally distributed technologies.

[21]The result that (7) holds under independence is very important for at least two reasons. First, independence is a standard assumption in models of growth and trade with multiple countries. Second, while this assumption is not appealing for empirical purposes, it is still very useful for developing the theory in the case of correlated distributions. In fact, it is often possible to find simple joint transformations of the variables such that the transformed variables become independent. In Appendix D, we exploit precisely this property to show that (7) also holds for some families of multivariate distributions that allow for correlated marginals.

[22]To the extent that the international diffusion of ideas raises the correlation between country technologies, our findings are consistent with Rodríguez-Clare (2007), who shows that the gains from trade are smaller when diffusion is included in the EK model.

[23]Introduced by Tawn (1990), $\Psi_{i,n}$ is also known as asymmetric bivariate logistic distribution and is com-

of independence" and is inversely related to the correlation between Z_i and Z_n: if $r = 1$, then Z_i and Z_n are independent (the case examined above); if $r < 1$, then Z_i and Z_n are positively correlated. As r goes to 0, the correlation between Z_i and Z_n tends to 1; in this case, we know from standard Ricardian theory that there are no comparative advantages to exploit and, therefore, both countries produce exactly as in autarky. Using (8), we can show that the TFP gain of country i, i.e. the increase in its TFP with respect to autarky, is:[24]

$$\frac{E\left(Z_i|Z_i \geq Z_n\right)}{E\left(Z_i\right)} = \left[1 + \left(\frac{T_n}{T_i}\right)^{1/r}\right]^{r/\theta} . \tag{9}$$

Hence, $E\left(Z_i|Z_i \geq Z_n\right) \geq E\left(Z_i\right)$ always.[25] Let us analyze this gain in two separate cases: $T_i = T_n$ and $T_i \neq T_n$.

Figure 1 shows, for $T_i/T_n = 1$, the TFP gain of country i for different values of θ and r.[26] We know from EK that welfare gains from trade are decreasing in θ; the figure shows that the same applies to TFP gains. With independent distributions, the TFP gain from trading with a symmetric country, i.e. one that has the same state of technology, goes from 7 percent (with $\theta = 10$) to 19 percent (with $\theta = 4$). In addition, for any value of θ the TFP gain is monotonically decreasing in the correlation between the technologies of i and n (or increasing in r).

Figure 2 shows the TFP gain for different values of r and T_i/T_n, given $\theta = 6.67$ (our reference value in the next section). Not surprisingly, the TFP gain is larger, the higher the productivity of the competitor country. If, for example, $T_i/T_n = 0.5$ the TFP gain for country i is as high as 18 percent with independent distributions, and goes down to 0 very slowly as r decreases; with $r = 0.1$ the TFP gain is still 11 percent. On the other hand, if $T_i/T_n = 2$ the TFP gain is at most 7 percent, and goes to zero more rapidly as r tends to zero. As before, the TFP gain decreases as correlation increases.

In order to show that the previous results are not specific to the multivariate Fréchet, but also hold for other families of distributions, we briefly illustrate TFP gains with normally distributed technologies.[27] Appendix D, provides other examples based on bivariate

---

monly used in multivariate extreme value theory.

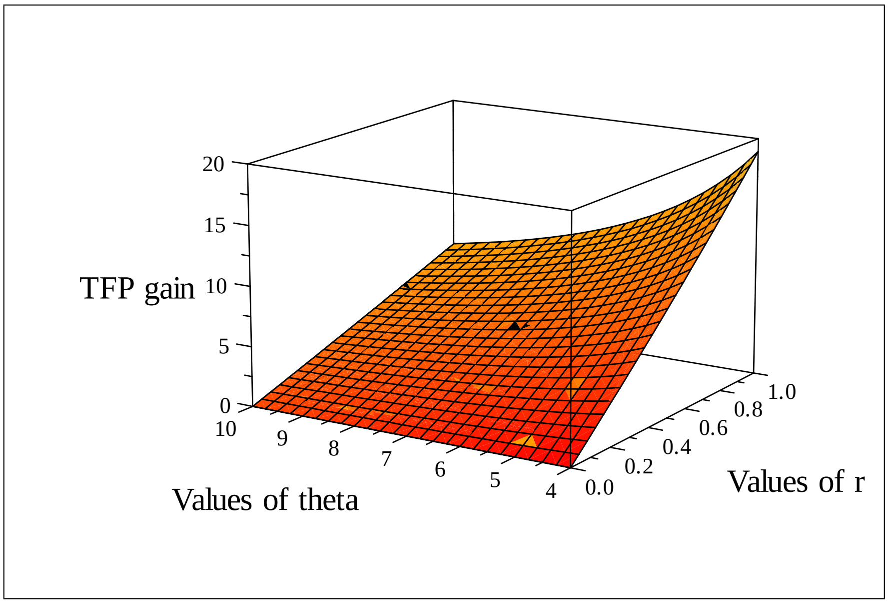
[24] See Appendix D, for details.

[25] An inspection of equation (9) reveals the important property that the TFP gain for two countries with correlated technologies and given values of T_i/T_n, θ and r (with $r < 1$) is the same as the TFP gain for two countries with independent technologies, a state-of-technology ratio equal to $\left(T_i/T_n\right)^{1/r}$, and a precision parameter equal to θ/r. Thus, we do not need to generalize Propositions 1 and 2 to the case of correlated Fréchet distributions: one can simply use the TFP gains derived under independence and obtain those under positive correlation with an appropriate rescaling of the parameters.

[26] Section 6 explains why we have placed θ in a range between 4 and 10.

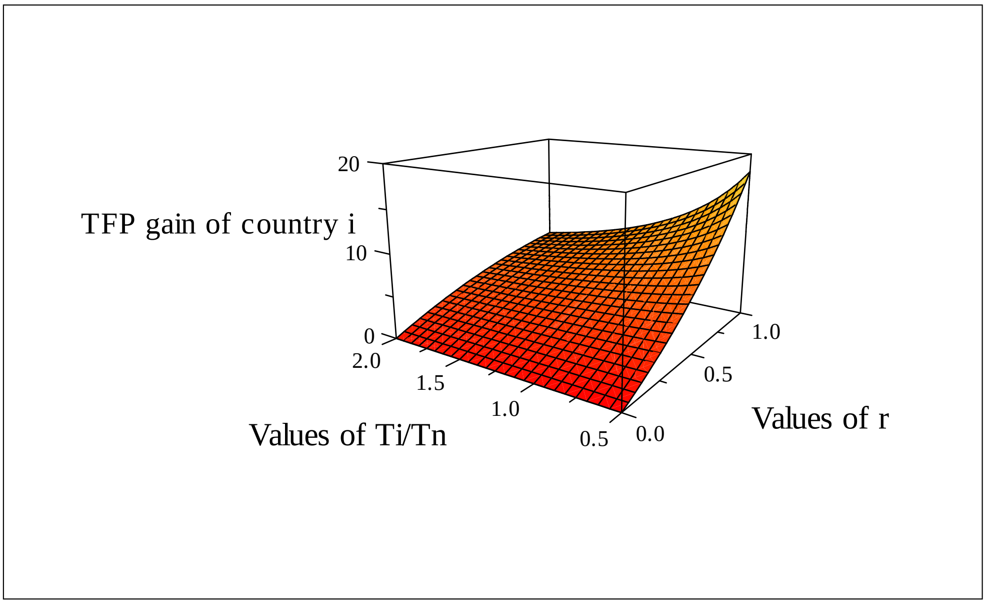
[27] This is quite a different case with respect to the Fréchet, given that the normal is a symmetric light-tailed distribution, that also allows to deal easily with negatively correlated technologies.

Figure 1: TFP gains from trade with a symmetric country (1)



(1) TFP gains from trade with respect to autarky, in percentages, for different values of r and θ, with T_i/T_n= 1.

Figure 2: TFP gains from trade with an asymmetric country (1)



(1) TFP gains from trade with respect to autarky, in percentages, for different values of r and T_i/T_n, with $\theta = 6.67$.

Pareto and lognormal distributions, as well as detailed computations for the normal case. For all these distributions, we can prove the result that $Z_i|Z_i \geq Z_n$ first-order stochastically dominates Z_i (that implies inequality (7)).

Thus, suppose that (Z_1, Z_2) has a bivariate normal distribution with the mean and variance of Z_i respectively denoted by μ_i and σ_i^2 $(i = 1, 2)$, and correlation ρ $(|\rho| < 1)$. For simplicity, assume $\sigma_1^2 = \sigma_2^2 = s^2$. The TFP gain of country 1 is:

$$\frac{E\left(Z_1|Z_1 \geq Z_2\right)}{E\left(Z_1\right)} = 1 + \frac{\sigma_v}{2\mu_1} \frac{g\left(\frac{\mu_2-\mu_1}{\sigma_v}\right)}{1 - G\left(\frac{\mu_2-\mu_1}{\sigma_v}\right)} \ , \tag{10}$$

where g and G are, respectively, the probability density function (p.d.f.) and the c.d.f. of the standard normal variable, and where $\sigma_v = s\sqrt{2\left(1-\rho\right)}$. Noting that the ratio $g/\left(1-G\right)$ is the hazard function of the normal distribution (which is strictly increasing in its own argument), it is easy to verify that all the main results obtained with the multivariate Fréchet are confirmed. Specifically, the TFP gain from trade of a country is always: non-negative; strictly increasing in the autarky TFP of the competitor country (μ_2), and in the degree of heterogeneity of domestic and foreign production (s); strictly decreasing in the domestic autarky TFP (μ_1) and in the correlation between domestic and foreign technologies (ρ). In particular, for what concerns the correlation, note that the largest TFP gain occurs as ρ tends to -1.

5 Weighted TFP

Throughout the paper, we define TFP as the first moment of the distribution of industry productivities. With this definition, then, industry productivities are not weighted. In this section, we consider a different definition of TFP in which we weight industry productivities with the value of industry production. This is a more correct theoretical counterpart of any empirical measure of TFP. Moreover, it is a definition in which demand fully kicks in and determines not only which industries produce some positive output, but also how much they produce.

Let us start from the closed economy and, for the sake of simplicity, let us neglect intermediate goods ($\beta = 1$) and non tradeables. By maintaining the assumption of CES preferences (with elasticity σ), demand for good j in country i is:

$$c_i\left(j\right) = \left[\frac{p_i}{p_i\left(j\right)}\right]^{\sigma} \frac{Y_i}{p_i} \ , \tag{11}$$

where Y_i denotes income, which is given by $w_i L_i$, with L_i being the number of workers. Hence, let us define the weighted TFP as:

$$\text{TFP}_i^w = \int_0^{+\infty} z_i \cdot \omega_i\left(z_i\right) dF_i\left(z_i\right) \ , \tag{12}$$

where we suppress the goods indices j in the integrals, and where the weight of good j is:

$$\omega_i (z_i (j)) = \frac{p_i (j) c_i (j)}{Y_i} = \frac{L_i (j)}{L_i} = \left[\frac{p_i}{p_i (j)}\right]^{\sigma-1} . \tag{13}$$

We can show the following result:

Proposition 3 *For any distribution Z_i such that the moment of order σ exists, we have:*

$$\mathrm{TFP}_i^w = \frac{E (Z_i^\sigma)}{E \left(Z_i^{\sigma-1}\right)} . \tag{P3}$$

Proof. See Appendix E.1 ■

In other words, in the closed economy the weighted TFP is equal to the ratio between the moment of order σ and the moment of order $\sigma - 1$ of the productivity distribution.

If we impose that Z_i is Fréchet distributed, with $\theta > \sigma$, so that the moment of order σ exists (see footnote 7), we immediately obtain that:

$$\mathrm{TFP}_i^w = E (Z_i) \cdot c (\sigma) , \tag{14}$$

with:

$$c (\sigma) = \frac{\theta}{\theta - \sigma} \left[B \left(\frac{\theta - \sigma + 1}{\theta}, \frac{\theta - 1}{\theta}\right)\right]^{-1} ,$$

where B denotes Euler's Beta function. Thus, with Fréchet distributed technologies, the weighted TFP is proportional to the first moment. The proportionality constant gathers the whole effect of demand. In particular, it is easy to show that $c(\sigma)$ and, therefore, TFP_i^w are increasing in σ. The intuition behind this result is straightforward. As the elasticity of substitution increases, consumers become more willing to substitute high- with low-price goods; then, demand for the latter increases. Low-price goods, in turn, are produced with a high efficiency and, since their weight rises, so does the weighted TFP.

Notice, in particular, the contrast between simple and weighted TFPs, that capture two distinct phenomena. The former is unaffected by consumer preferences, as it only depends on "intrinsic" technology (that, in a Schumpeterian fashion, can be thought of as a variable determined by innovation and selection). The latter, on the contrary, is affected by consumer preferences, because substitutability of goods determines the demand for each good, the value of industry output and, in turn, the productivity weights.

It is important to note that the linear relationship between TFP and the first moment of the productivity distribution obtained in equation (14) is not specific to the Fréchet, but also holds for other commonly used distributions, such as the Pareto and the lognormal. In particular, if Z_i is $Pareto(\xi, \alpha)$ (i.e. $F_i(z) = 1 - (\xi/z)^\alpha$ with $\alpha > \sigma$), then equation (14) still holds with $c(\sigma) = (\alpha - 1)(\alpha - \sigma + 1)/\alpha(\alpha - \sigma)$. Analogously, if Z_i is

Lognormal (μ, s^2) (i.e. $Z_i = \exp(X_i)$ and $X_i \sim Normal(\mu, s^2)$), then equation (14) holds with $c(\sigma) = \exp\left[s^2(\sigma - 1)\right]$. For both distributions, $c(\sigma)$ is still increasing in the elasticity of substitution σ.

What about other distributions? If we assume that $\sigma \in \mathbb{N}$ (still maintaining that the moment of order σ exists), the following result holds:

$$E(Z_i^\sigma) = [E(Z_i)]^\sigma + \sum_{k=0}^{\sigma-2} \binom{\sigma}{k} \beta_{\sigma-k} [E(Z_i)]^k \ ,$$

where $\beta_{\sigma-k}$ is the central moment of order $\sigma - k$ of Z_i (see Balakrishnan and Nevzorov, 2003). Thus, for any random variable Z_i, its moment of order σ is a polynomial of degree σ in the first moment $E(Z_i)$, with coefficients depending on the central moments of Z_i. It follows that:

$$\text{TFP}_i^w = E(Z_i) \cdot \left(1 + \frac{1}{O\left\{[E(Z_i)]^2\right\}}\right) \ .$$

Hence, for any distribution Z_i with $E(Z_i)$ sufficiently large, we can measure the weighted TFP with the first moment. The largest measurement error that we can make converges to zero as quickly as $1/[E(Z_i)]^2$.

Let us turn to the open economy and consider, first, what happens when the country moves from autarky to no trade barriers (i.e. $d_{ni} = 1$ for any n and i). One can easily check that, for any autarky distribution of productivities, the weighted TFP of the open economy is still given by the ratio between the moment of order σ and the moment of order $\sigma - 1$ of the productivity distribution of the surviving industries. In other words, Proposition 3 still holds with Z_i (the autarky distribution) replaced by $Z_{i,o}$ (the distribution for the industries that survive international competition).

If we impose the Fréchet assumption, the ratio between the weighted TFP of the open and the closed economy is then given by the ratio between the first moments of $Z_{i,o}$ and Z_i — i.e. exactly the same ratio that we have considered in the previous sections. In fact, since $Z_{i,o}$ has the same precision parameter as Z_i, the proportionality constant $c(\sigma)$ is the same in the open and the closed economy and cancels out when we take the ratio between the two TFPs. Hence, in the case of "free trade", considering the weighted instead of the non-weighted TFP would not alter our main insights. Key to this result is that foreign consumers have the same CES preferences as domestic consumers. Therefore, they demand the good j produced by country i proportionally to $z_i^{\sigma-1}(j)$, exactly as the domestic consumers of country i do. The extra demand coming from abroad does not modify the composition of the demand for domestic goods. Thus, if consumer preferences are identical across countries, then when any country moves from autarky to free trade its aggregate TFP is affected only by a selection effect and consumer preferences do not play any role.

With finite trade barriers, instead, consumer preferences do play a role, even when they are identical all over the world. A simple example with two countries (n and i), no intermediate goods ($\beta = 1$), no tradeable goods (so that $Y_i = w_i L_i$) and Fréchet distributed technologies is sufficient to build some intuition. In Appendix E.1 we prove the following result:

$$\mathrm{TFP}^{w}_{i,o} = \frac{E\left(Z^{\sigma}_{i,o}\right)}{E\left(Z^{\sigma-1}_{i,o}\right)} + b\frac{E\left(Z^{\sigma}_{i,e}\right)}{E\left(Z^{\sigma-1}_{i,o}\right)} \ , \tag{15}$$

where $Z_{i,e}$ is the random variable that describes the productivities of the exporters (which is a Fréchet with state $T_i + T_n \left(w_i d_{ni}/w_n\right)^{\theta}$ and precision θ), and $b \geq 0$ is a constant term that depends on the "relative weight" of country n with respect to country i (measured by real income) as well as on other fundamentals, including trade barriers.[28] Equation (15) shows that two distinct effects are at work. The weighted TFP of the open economy is still raised by the exposure to international competition that forces less efficient firms to shut down — *a "pure" selection effect*. This effect is expressed by the first addendum of the right-hand side of equation (15). As before, with Fréchet distributed technologies, this effect is measured by the ratio between the first moments of $Z_{i,o}$ and Z_i (the ratio between the moment of order σ and the moment of order $\sigma - 1$ simplifies into the first moment, for both $Z_{i,o}$ and Z_i). However, now there is also a second effect (the second addendum of the right-hand side of equation (15)), that depends on the productivity of exporters. After opening to trade, in fact, exporters produce relatively more than non-exporters with respect to what they did under autarky, because only these producers benefit from some extra demand coming from abroad.[29] Thus, the additional foreign demand for the domestic goods that are exported leads to an increase in the weight of exported versus non-exported goods — *a market-share reallocation effect*.[30] Since, on average, exporters are more productive than non-exporters (as shown by the higher state of technology of $Z_{i,e}$), this effect further raises the weighted TFP.[31]

[28] Appendix E.1 also shows that as the country approaches "free trade" (i.e. as $d_{ni} \to 1$), then $\mathrm{TFP}^{w}_{i,o}$ goes to $E\left(Z^{\sigma}_{i,o}\right)/E\left(Z^{\sigma-1}_{i,o}\right)$ that, in turn, becomes equal to $E\left(Z^{\sigma}_{i,e}\right)/E\left(Z^{\sigma-1}_{i,e}\right)$. On the other hand, as the country approaches autarky (i.e. as $d_{ni} \to +\infty$), then b goes to zero and $Z_{i,o} \to Z_i$.

[29] Let j' (j'') be a good that, after opening to trade, survives (does not survive) international competition and is (is not) exported. Under autarky, the relative expenditure of domestic consumers on good j' with respect to good j'' is equal to $z^{\sigma-1}_{i}\left(j'\right)/z^{\sigma-1}_{i}\left(j''\right)$ (see Appendix E.1). After opening to trade, the relative expenditure of domestic consumers is still given by $z^{\sigma-1}_{i}\left(j'\right)/z^{\sigma-1}_{i}\left(j''\right)$ but, now, the is also some extra expenditure on j' by foreign consumers.

[30] This market-share reallocation effect is consistent with the evidence presented, e.g., in Pavcnik (2002) and Bernard, Eaton, Jensen, and Kortum (2003).

[31] These results yield some implications for the effect of trade on the TFP of the whole economy. If we compute it as a non-weighted average of all (tradeable and non-tradeable) industries, then trade always raises TFP because the TFP of tradeables (weighted or non-weighted) rises after opening while that of the non-tradeables does not change. If we measure the TFP of the whole economy with a weighted average of industry productivities, the TFP of the whole economy certainly rises if labor is immobile between the tradeable and the non-tradeable sector. If it is mobile, instead, a share of labor may flow from or to the non-tradeable sector, and the effect on the weighted TFP of the whole economy will depend on the relative productivities of tradeable and non tradeable industries.

In the general setting with many countries, working out an analytic expression for the market-share reallocation effect is extremely cumbersome (see Appendix E.2). Therefore, in the next section we stick to the "pure" selection effect. This effect is a lower bound to the overall effect of international competition. It is a good approximation of the overall effect of trade if trade barriers are negligible or if the domestic country is large enough with respect to the foreign country (in this case, b is very small).

6 Quantifying the selection effect

Let us go back to the baseline model described in Sections 2 and 3. An immediate implication of Propositions 1 and 2 is that the contribution of international competition to the TFP of the tradeable sector (hereafter identified with the manufacturing sector) due to the section effect is simply given by the measure of openness Ω_i raised to the $1/\theta$ power (one can obtain it by substituting (P2) into (5) and dividing the result by the mean of Z_i). In this section we quantify this effect for a sample of 41 countries for which we have the data on manufacturing production and trade required to measure Ω_i. We carry out the exercise for the period 1985-2005.[32]

A value for θ can be obtained using two different approaches. One is followed by EK, who estimate θ using some testable implications of the model and find values between 3 and 13 (their preferred estimate is 8.28). An alternative strategy is proposed by Alvarez and Lucas (2007), who calibrate θ by exploiting a connection between the models of EK and Armington (1969). Although in the latter goods produced in different countries are treated as different, the prediction that market shares are given by equation (1) also obtains in the Armington model by replacing θ with $\sigma_a - 1$, where σ_a is the Armington elasticity. Based on the estimates of import elasticities surveyed by Anderson and van Wincoop (2004), Alvarez and Lucas consider 6.67 as their benchmark calibration, from a range between 4 and 10. We choose to follow Alvarez and Lucas and set $\theta = 6.67$.[33]

Table 1 shows the contributions of international competition to the TFP of our sample countries at five-year intervals between 1985 and 2005. In 2005, international competition lifts the manufacturing TFP of the equally-weighted average country by 11 percent above its autarky level (9 percent when calculated on the median). Across countries, the gain from international competition ranges from 2 percent for India, to 39 percent for the Netherlands.[34]

[32] For a detailed description of data sources and the empirical methodology, see Appendix F.

[33] Notice that both Alvarez and Lucas and EK consider cross-sectional data. In our time-series analysis, we take θ time-invariant. Finicelli, Pagano, and Sbracia (2009b) provide some evidence supporting this assumption.

[34] The estimates of $\Omega_i^{1/\theta}$ for different values of θ can be derived with simple back-of-the-envelope calculations. Setting $\theta = 8.28$ (the preferred estimate of EK), in particular, the values reported in Table 1 would be slightly

This contribution tends to be smaller for larger countries; its negative correlation with the country size, however, is far from being perfect.[35]

Over time, the average contribution of international competition exhibits a neat positive trend (from 6 to 11 percent between 1985 and 2005), which is shared by most countries. For 9 countries, however, the contribution of international competition does not pick in the latest available year. Notice also that between the years 1995 and 2000, a period that includes the Asian and Russian crises, the contribution of international competition declines for 7 out of the 37 countries for which we have data for both years, with a particularly large drop for Malaysia and Singapore (of 6 percentage points).

With some simple algebra, one can derive from the theory also the effect of international competition on the real wage in the manufacturing sector (w_i/p_i), a measure of welfare in EK, and find that this is equal to Ω_i raised to the $1/\beta\theta$ power (see also equation (15) in EK). By setting $\beta = 0.33$ — a calibration in line with those of the literature — and retaining $\theta = 6.67$, one can immediately obtain the values of $\Omega_i^{1/\beta\theta}$ from Table 1.[36] For the equally-weighted average country, we find that in 2005 international competition raises the real wage by over 35 percent with respect to the level that would have been observed under autarky (the rule of thumb is that, with $\beta = 0.33$, the effect on the real wage is approximately three times larger than the effect on TFP).

To understand this result, suppose that the nominal wage is constant (in other words, take the nominal wage as the numéraire). Then, the price level in the manufacturing sector would decline by over 35 percent with respect to the closed economy. Because of perfect competition, here the entire productivity gain is translated into lower prices, an effect that accounts for 11 percentage points. However, there is also an indirect effect stemming from the fact that a share $1 - \beta$ of the manufacturing goods also serves as intermediate goods. Hence, the TFP effect on the price level is amplified by the availability of lower-price intermediate inputs. In this example, with $\beta = 0.33$, the overall TFP effect is three times larger than the direct effect alone. (If only a subset of intermediate goods were tradeable, then this amplification effect would be smaller).

smaller. For instance, in 2005 the average gain across countries would be around 9 percent.

[35] By measuring the country size by population, GDP in PPP or current prices, manufacturing production, or manufacturing value added, we never obtain a correlation below -0.30.

[36] The parameter β can be calibrated using two different approaches. EK calibrate it as the cross-country average of the labor share in gross manufacturing production. This calibration implies that labor is the sole production factor and capital goods are comprised into intermediate goods. Alvarez and Lucas (2007), instead, calibrate β as the cross-country average of manufacturing value added over gross manufacturing production. By doing so, these authors consider labor plus capital goods as the single production factor, which they label as 'equipped labor'. For the 19 countries originally considered by EK, Finicelli, Pagano, and Sbracia (2008a) report that, over the period 1985-2002, the first (second) calibration would provide annual values of β between 0.19 (0.31) and 0.22 (0.34). Alvarez and Lucas also report that, using their calibration, the world average value of β in 1998 (from the UNIDO Industrial Statistics database) would be equal to 0.38.

Table 1: Contribution of the selection effect to TFP in selected years (1)

country	1985	1990	1995	2000	2005
Argentina*	0.9 [2]	2.2	2.9	4.3	4.9
Australia	3.4	3.5	5.0	7.0	4.4
Austria	8.7	10.4	11.7	16.2	18.3
Belgium*			9.5	11.1	14.0 [7]
Brazil*				2.9	3.0 [7]
Canada	6.9	7.2	10.8	11.8	10.6
Chile*		4.9	5.8	5.4	4.8
Czech Republic			6.7	11.5	13.4
Denmark	10.6	11.6	13.0	17.1	21.9
Finland	5.2	5.6	6.8	7.9	8.5
France	4.4	5.3	6.1	7.6	8.3
Germany	5.2	5.2	5.8	8.0	9.3
Greece	3.9	5.6	6.3	8.4	7.6
Hungary		6.6	8.1	16.4	18.6
Iceland	13.0	13.9	15.1	18.4	24.8
India*	1.2	1.1	1.6	1.9	2.4 [7]
Indonesia*	5.2	6.6	6.4	5.6	4.9
Ireland	13.5	15.2	15.8	18.2	20.0 [5]
Israel*			9.4	9.2	9.8 [7]
Italy	3.7	3.6	4.7	5.4	5.6
Japan	1.6	1.5	1.7	2.2	2.7
Republic of Korea	3.5	3.5	4.2	4.8	4.8
Malaysia*	8.7	10.0	14.7	9.0	9.2 [6]
Mexico		2.1	7.3	9.5	10.1
Netherlands	14.9	17.1	16.8	26.4	39.3
New Zealand	6.3	6.9	7.3	8.9	
Norway	9.0	9.6	9.7	11.2	10.9
Philippines*	3.4	4.4	6.2		
Poland		3.4	4.7	7.0	9.3
Portugal	3.3	6.1	7.2	9.9	10.2
Russian Federation*			3.7 [3]	3.2	4.1
Singapore*	16.2	17.0	22.4	16.6	23.6 [7]
Slovak Republic				14.6	21.1
South Africa*		2.1	4.0	3.0	3.4
Spain	2.7	3.5	4.5	6.3	6.6
Sweden	6.8	6.7	8.5	9.7	11.1
Switzerland		8.8	9.3	14.5	15.6
Thailand*	2.2 [2]	3.5	6.0 [4]		
Turkey*		3.2	4.4	5.8	
United Kingdom	5.6	6.1	7.6	9.1	10.1 [6]
United States	1.6	2.2	2.7	3.5	3.9
mean	6.1	6.5	7.8	9.5	11.1
median	5.2	5.6	6.7	8.9	9.3
number of countries	28	35	39	39	37

Source: authors' calculations on OECD STAN data; * UNIDO IDSB data.

Notes: (1) Values of $\Omega_i^{1/\theta}-1$, in percentage, for each country i; (2) 1984; (3) 1994; (4) 1996; (5) 2002; (6) 2003; (7) 2004.

Thanks to the amplification effect on real wages, the welfare gains are non-negligible, even for large countries where the mere TFP gains due to the selection effect reported in Table 1 appear small. For instance, suppose that the United States raise trade barriers so as to go back to the degree of openness observed in 1985. This would imply a decline of 2.3 percent in their TFP, and a drop as large as 7.3 percent in their real wages.

7 Conclusion

Exploiting the probabilistic formulation of the Ricardian model developed by Eaton and Kortum (2002), we have analyzed the theoretical foundations of the relationship between trade and TFP. We have shown that the correlation between the autarky distributions of industry productivities is key in determining the sign and size of TFP growth after opening to trade. First, we have established that independence yields the remarkable implication that trade openness always raises TFP. Second, if technologies are correlated this finding, while not generalizable, still holds for important families of joint distributions of productivities, such as the multivariate Fréchet, Pareto, normal, and lognormal. These results warrant further research on the statistical distribution of productivities (a novel primitive of modern models of growth and international trade), an issue about which the literature is still inconclusive.

Consistently with the Ricardian approach, the law of comparative advantage is the driving force behind the selection effect. In this setting, any industry, even a high-productivity one, can exit the market, but this happens with a probability that is lower for higher-productivity industries.

Our analysis also delivers a simple model-based measure of the selection effect that has a straightforward empirical implementation. We show that the TFP of an open economy is equal to its TFP under autarky augmented by a factor that depends solely on its aggregate production and trade. This finding, which is unaffected by measuring TFP with a weighted instead of a non-weighted average of industry productivities, enables to easily quantify the magnitude of TFP gains from trade.

Appendix

A Proof of Proposition 1

Before computing $F_{i,o}$ from equation (4), we show that the goods produced by country i are all and only those for which it holds $p_{ii}(j) \leq p_{ik}(j)$ for any k. If $p_{ii}(j) \leq p_{ik}(j)$ for any k, then good j is produced by country i and sold at home. Hence, we only need to show that there is no good j which is produced by country i, exported in a country $n \neq i$, and not sold at home. Clearly, if such a good is not sold at home, it means that there is another country, call it k ($k \neq i$), that sells it in country i at a lower cost. More formally, then, we need to show that there is no good j such that: (i) $p_{ii}(j) > p_{ik}(j)$ for some k; and (ii) $p_{ni}(j) < p_{nl}(j)$ for some n and for any $l \neq i$. Suppose, by contradiction, that there exists such a good j. The inequality (i) means that: $c_i/z_i(j) > c_k d_{ik}/z_k(j)$. The inequality (ii) is equivalent to: $c_i d_{ni}/z_i(j) < c_l d_{nl}/z_l(j)$ for any $l \neq i$. Now take $l = k$. Then: $c_i d_{ni}/z_i(j) < c_k d_{nk}/z_k(j)$. However, from the first inequality we can also obtain: $c_i d_{ni}/z_i(j) > c_k d_{ik} d_{ni}/z_k(j) \geq c_k d_{nk}/z_k(j)$, where the last part follows from the triangle inequality and contradicts the inequality (ii).

We now turn to the computation of $F_{i,o}(z)$. To find the distribution of the TFP of country i (TFP_i), we consider first the price distribution of the goods that country i "submits" to country n. Denote this random variable by P_{ni} and its c.d.f. by W_{ni}. Recalling that $p_{ni}(j) = c_i d_{ni}/z_i(j)$ for any good j, EK show that:

$$W_{ni}(p) = \Pr(P_{ni} \leq p) = 1 - F_i\left(\frac{c_i d_{ni}}{p}\right) = 1 - \exp\left[-T_i (c_i d_{ni})^{-\theta} p^{\theta}\right] ,$$

where F_i is the c.d.f. of Z_i. By setting: $\phi_{ni} = T_i (c_i d_{ni})^{-\theta}$, we can write the p.d.f. of P_{ni} as:

$$w_{ni}(p) = \phi_{ni} \cdot \theta \cdot p^{\theta-1} \cdot \exp\left(-\phi_{ni} \cdot p^{\theta}\right) ;$$

thus, P_{ni} has a Weibull distribution.

Now let us turn to TFP_i, whose distribution is:

$$F_{i,o}(z) = \Pr\left(Z_i < z | P_{ii} = \min_k P_{ik}\right) = \frac{\Pr\left(P_{ii} = \min_k P_{ik} \ , \ Z_i < z\right)}{\Pr\left(P_{ii} = \min_k P_{ik}\right)} .$$

The denominator corresponds to equation (8) of EK for $n = i$; namely:

$$\Pr\left(P_{ii} = \min_k P_{ik}\right) = \Pr(P_{ii} \leq P_{i1}, ..., P_{ii} \leq P_{iN}) = \frac{T_i c_i^{-\theta}}{\sum_{k=1}^{N} T_k (c_k d_{ik})^{-\theta}} .$$

The numerator is:

$$\begin{aligned}
\Pr\left(P_{ii}=\min_k P_{ik}\ ,\ Z_i<z\right) &= \Pr\left(P_{ii}\le P_{i1},...,P_{ii}\le P_{iN},Z_i<z\right)= \\
&= \Pr\left(Z_1\le \frac{Z_i c_1 d_{i1}}{c_i},...,Z_N\le \frac{Z_i c_N d_{iN}}{c_i},\ Z_i<z\right)= \\
&= \int_0^z \prod_{k\neq i} F_k\left(\frac{z_i c_k d_{ik}}{c_i}\right)\cdot f_i\left(z_i\right)dz_i= \\
&= \frac{T_i c_i^{-\theta}}{\sum_{k=1}^{N} T_k\left(c_k d_{ik}\right)^{-\theta}}\cdot\int_0^z \Lambda_i\cdot\theta\cdot z_i^{-(\theta+1)}\cdot\exp\left(-\Lambda_i\cdot z_i^{-\theta}\right)dz_i\ ,
\end{aligned}$$

where Λ_i is given by equation (P1).

By using the expressions found for the numerator and the denominator of $F_{i,o}(z)$, we have that:

$$F_{i,o}(z)=\int_0^z \Lambda_i\cdot\theta\cdot x^{-(\theta+1)}\cdot\exp\left(-\Lambda_i\cdot x^{-\theta}\right)dx\ ;$$

in other words, $Z_{i,o}\sim \textit{Fréchet}\left(\Lambda_i,\theta\right)$.

B Proof of Proposition 2

Plugging the expression of costs (equation (3)) into equation (P1), and multiplying and dividing by T_i we can write:

$$\Lambda_i=T_i+T_i\sum_{k\neq i}\frac{T_k}{T_i}\left(\frac{w_k}{w_i}\right)^{-\theta\beta}\left(\frac{p_k}{p_i}\right)^{-\theta(1-\beta)}d_{ik}^{-\theta}\ .$$

Using equation (1), we can obtain:

$$\frac{X_{ik}}{X_{ii}}=\frac{X_{ik}/X_i}{X_{ii}/X_i}=\frac{T_k}{T_i}\left(\frac{w_k}{w_i}\right)^{-\theta\beta}\left(\frac{p_k}{p_i}\right)^{-\theta(1-\beta)}d_{ik}^{-\theta}\ .$$

Therefore, substituting back into Λ_i we find:

$$\Lambda_i=T_i\left(1+\sum_{k\neq i}\frac{X_{ik}}{X_{ii}}\right)\ .$$

C Conditional mean under independence

In this section we prove that $E\left(Z_i|Z_i\ge Z_n\right)\ge E\left(Z_i\right)$ for any Z_i and Z_n independent random variables (which we take with absolutely continuous distributions and support in $\mathbb{R}$) with, for obvious reasons, $\Pr(Z_i \geq Z_n) > 0$. Let us denote with f_i and f_n the p.d.f. of Z_i and Z_n; $f_{i,n} = f_i \cdot f_n$ denotes the p.d.f. of the random vector (Z_i, Z_n).

We can write:

$$\begin{aligned}
E(Z_i | Z_i \geq Z_n) &= \frac{1}{\Pr(Z_i \geq Z_n)} \int_{\mathbb{R}} z_i \left[\int_{z_i \geq z_n} f_{i,n}(z_i, z_n)\, dz_n \right] dz_i \\
&= \frac{1}{\Pr(Z_i \geq Z_n)} \int_{\mathbb{R}} \int_{\mathbb{R}} z_i \cdot f_{i,n}(z_i, z_n) \cdot I_S(z_i, z_n)\, dz_i dz_n \ ,
\end{aligned}$$

where I_S denotes the indicator function of the set S and:

$$S = \left\{ (z_i, z_n) \in \mathbb{R}^2 : z_i \geq z_n \right\} \ .$$

Hence, $E(Z_i | Z_i \geq Z_n) \geq E(Z_i)$ if and only if:

$$\int_{\mathbb{R}} \int_{\mathbb{R}} z_i \cdot f_{i,n}(z_i, z_n) \cdot I_S(z_i, z_n)\, dz_i dz_n \geq E(Z_i) \cdot \Pr(Z_i \geq Z_n) \ . \tag{16}$$

We now show the intermediate result that:

$$\int_{\mathbb{R}} z_i \cdot f_i(z_i) \cdot I_{[z_n, +\infty)}(z_i)\, dz_i \geq E(Z_i) \cdot \Pr(Z_i \geq z_n) \ , \forall z_n \in \mathbb{R}. \tag{17}$$

There are three possible cases: (i) if $\Pr(Z_i \geq z_n) = 0$, then both sides of the inequality are equal to zero; (ii) if $\Pr(Z_i \geq z_n) = 1$, then both sides of the inequality are equal to $E(Z_i)$; (iii) if $0 < \Pr(Z_i \geq z_n) < 1$, then from:

$$E(Z_i | Z_i \geq z_n) \geq z_n \geq E(Z_i | Z_i < z_n) \ ,$$

it follows that:

$$\frac{\int_{\mathbb{R}} z_i \cdot f_i(z_i) \cdot I_{[z_n, +\infty)}(z_i)\, dz_i}{\Pr(Z_i \geq z_n)} \geq \frac{\int_{\mathbb{R}} z_i \cdot f_i(z_i) \cdot I_{(-\infty, z_n)}(z_i)\, dz_i}{\Pr(Z_i < z_n)} \ .$$

Multiplying both sides of the inequality by $\Pr(Z_i \geq z_n) \cdot \Pr(Z_i < z_n)$ and adding them the term:

$$\Pr(Z_i \geq z_n) \cdot \int_{\mathbb{R}} z_i \cdot f_i(z_i) \cdot I_{[z_n, +\infty)}(z_i)\, dz_i \ ,$$

completes the proof that (17) holds.

We can now integrate the inequality (17) with respect to the p.d.f. of Z_n:

$$\int_{\mathbb{R}} \left[\int_{\mathbb{R}} z_i \cdot f_i(z_i) \cdot I_{[z_n, +\infty)}(z_i)\, dz_i \right] \cdot f_n(z_n)\, dz_n \geq E(Z_i) \int_{\mathbb{R}} \Pr(Z_i \geq z_n) \cdot f_n(z_n)\, dz_n \ .$$

We finally use the independence assumption. First, for what concern the right-hand side, we have that:

$$\begin{aligned}
\int_{\mathbb{R}} \Pr(Z_i \geq z_n) f_n(z_n)\, dz_n &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f_i(z_i) I_{[z_n, +\infty)}(z_i)\, dz_i \right] \cdot f_n(z_n)\, dz_n \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} f_{i,n}(z_i, z_n) \cdot I_S(z_i, z_n)\, dz_i dz_n \\
&= \Pr(Z_i \geq Z_n) \ .
\end{aligned}$$

Using again the independence assumption, we can rewrite the left-hand side as:

$$\int_{\mathbb{R}} \int_{\mathbb{R}} z_i \cdot f_{i,n}\left(z_i, z_n\right) \cdot I_S\left(z_i, z_n\right) dz_i dz_n \ ,$$

and this proves the necessary and sufficient condition (16).

D TFP gains with correlated distributions

In this section we compute TFP gains for a variety of multivariate distributions. We focus on common distributions whose margins are consistent with those considered by theoretical or empirical studies about the distribution of plant-, firm-, or industry-level productivities or sizes, such as the Fréchet, Pareto, and lognormal distributions. Notice that these distributions are right-skewed and heavy-tailed. Thus, we also consider the normal distribution in order to confirm that the result that TFP increases after opening to trade also holds for symmetric distributions with light (exponential) tails. For all these distributions, we prove the result that $Z_i|Z_i \geq Z_n$ first-order stochastically dominates Z_i, a result that implies: $E\left(Z_i|Z_i \geq Z_n\right) \geq E\left(Z_i\right)$.

D.1 Fréchet

Recall that, for any $z_i > 0$ and $z_n > 0$:

$$\Psi_{i,n}\left(z_i, z_n\right) = \exp\left\{-\left[\left(T_i \cdot z_i^{-\theta}\right)^{1/r} + \left(T_n \cdot z_n^{-\theta}\right)^{1/r}\right]^r\right\} .$$

is the c.d.f. of a bivariate Fréchet. The corresponding p.d.f., which we denote with $\psi_{i,n}$, is a complicated function of (z_i, z_n). However, notice that the distribution of $Z_i|Z_i \geq Z_n$ is:

$$\begin{aligned} \Pr\left(Z_i < z|Z_i \geq Z_n\right) &= \frac{\Pr\left(Z_i < z, Z_i \geq Z_n\right)}{\Pr\left(Z_i \geq Z_n\right)} \\ &= \frac{\int_0^z \left[\int_0^{z_i} \psi_{i,n}(z_i, z_n) dz_n\right] dz_i}{\int_0^{+\infty} \left[\int_0^{z_i} \psi_{i,n}(z_i, z_n) dz_n\right] dz_i} . \end{aligned} \tag{18}$$

In order to compute it, we need to calculate:

$$\begin{aligned} \int_0^{z_i} \psi_{i,n}(z_i, z_n) dz_n &= \int_0^{z_i} \frac{\partial}{\partial z_n}\left[\frac{\partial}{\partial z_i}\Psi_{i,n}\left(z_i, z_n\right)\right] dz_n \\ &= \left[\frac{\partial}{\partial z_i}\Psi_{i,n}\left(z_i, z_n\right)\right]_{z_n \to z_i} - \left[\frac{\partial}{\partial z_i}\Psi_{i,n}\left(z_i, z_n\right)\right]_{z_n \to 0} . \end{aligned} \tag{19}$$

Therefore, we only have to compute the first derivative of $\Psi_{i,n}$ with respect to z_i, which is equal to:

$$\frac{\theta}{z_i}\left(\frac{T_i}{z_i^\theta}\right)^{\frac{1}{r}}\left(\left(\frac{T_i}{z_i^\theta}\right)^{\frac{1}{r}} + \left(\frac{T_n}{z_n^\theta}\right)^{\frac{1}{r}}\right)^{r-1} \exp\left(-\left(\left(\frac{T_i}{z_i^\theta}\right)^{\frac{1}{r}} + \left(\frac{T_n}{z_n^\theta}\right)^{\frac{1}{r}}\right)^r\right) .$$

It is easy to check that equation (19) is equal to:

$$\theta T_i^{\frac{1}{r}} z_i^{-\theta-1}\left(T_i^{\frac{1}{r}}+T_n^{\frac{1}{r}}\right)^{r-1}\exp\left(-\frac{1}{z_i^{\theta}}\left(T_i^{\frac{1}{r}}+T_n^{\frac{1}{r}}\right)^{r}\right) .$$

Hence, we can compute the denominator of (18):

$$\Pr\left(Z_i \geq Z_n\right)=\frac{T_i^{\frac{1}{r}}}{T_i^{\frac{1}{r}}+T_n^{\frac{1}{r}}} .$$

Computing the numerator of (18) and dividing it by $\Pr\left(Z_i \geq Z_n\right)$, we immediately recognize that:

$$Z_i|Z_i \geq Z_n \sim \textit{Fréchet}\left(\left(T_i^{\frac{1}{r}}+T_n^{\frac{1}{r}}\right)^{r},\theta\right) .$$

Hence, taking the first moment of this distribution and dividing it by the first moment of Z_i we find the TFP gain. Since the state of technology of $Z_i|Z_i \geq Z_n$ is larger than that of Z_i, the properties of the Fréchet distribution imply that $Z_i|Z_i \geq Z_n$ first-order stochastically dominates Z_i.

D.2 Pareto (Mardia's Type I)

For $z_i > 1$, $z_n > 1$, and $\lambda > 0$, the random vector (Z_i, Z_n) has a standard bivariate Pareto distribution of Mardia's Type I (see Mardia, 1970) if its p.d.f. is:

$$f_{i,n}\left(z_i,z_n\right)=\lambda\left(\lambda+1\right)\frac{1}{\left(z_i+z_n-1\right)^{(\lambda+2)}} .$$

The correlation between Z_i and Z_n is $1/\lambda$ (if $\lambda > 2$). We assume that $\lambda > 2$ to grant that first and second moments exist.

The p.d.f. of Z_i and Z_n are $f_i\left(z\right)=f_n\left(z\right)=\lambda z^{-\lambda-1}$, the corresponding c.d.f. are $F_i\left(z\right)=F_n\left(z\right)=1-z^{-\lambda}$, and their first moments are equal to:

$$E\left(Z_i\right)=E\left(Z_n\right)=\frac{\lambda}{\lambda-1},\text{ for }\lambda>1 .$$

We want to compute:

$$\begin{aligned}\Pr\left(Z_i < z|Z_i \geq Z_n\right) &= \frac{\Pr\left(Z_i < z, Z_i \geq Z_n\right)}{\Pr\left(Z_i \geq Z_n\right)}\\ &= \frac{\int_1^z\left[\int_1^{z_i} f_{i,n}(z_i,z_n)dz_n\right]dz_i}{\int_1^{+\infty}\left[\int_1^{z_i} f_{i,n}(z_i,z_n)dz_n\right]dz_i} .\end{aligned}\tag{20}$$

Hence, we need to calculate:

$$\begin{aligned}\int_1^{z_i} f_{i,n}(z_i,z_n)dz_n &= \lambda\left(\lambda+1\right)\int_1^{z_i}\frac{1}{\left(z_i+z_n-1\right)^{(\lambda+2)}}dz_n\\ &= \lambda\left[\frac{1}{z_i^{\lambda+1}}-\frac{1}{\left(2z_i-1\right)^{\lambda+1}}\right] .\end{aligned}$$

Computing the denominator of (20) we obtain:

$$\begin{aligned}\Pr\left(Z_i \geq Z_n\right) &= \int_1^{+\infty} \lambda \left[\frac{1}{z_i^{\lambda+1}} - \frac{1}{(2z_i-1)^{\lambda+1}}\right] dz_i \\ &= \lim_{z_i \to 1} \left[\frac{1}{z_i^{\lambda}} - \frac{1}{2}\frac{1}{(2z_i-1)^{\lambda}}\right] \\ &= \frac{1}{2} \ .\end{aligned}$$

Computing the numerator of (20) we obtain:

$$\int_1^{z} \lambda \left[\frac{1}{z_i^{\lambda+1}} - \frac{1}{(2z_i-1)^{\lambda+1}}\right] dz_i = \frac{1}{2}\frac{1}{(2z-1)^{\lambda}} - \frac{1}{z^{\lambda}} + \frac{1}{2} \ .$$

Hence, the c.d.f. of $Z_i | Z_i \geq Z_n$ is:

$$F\left(z\right) = \frac{1}{(2z-1)^{\lambda}} - \frac{2}{z^{\lambda}} + 1 \ ,$$

which has support in $[1, +\infty)$. Note that

$$F\left(z\right) \leq F_i\left(z\right) \ \forall z > 1 \ ;$$

in other words, $Z_i | Z_i \geq Z_n$ first-order stochastically dominates Z_i. It follows that $E\left(Z_i | Z_i \geq Z_n\right) \geq E\left(Z_i\right)$. In particular, the p.d.f. of $Z_i | Z_i \geq Z_n$ is:

$$f\left(z\right) = 2\lambda \left(\frac{1}{z^{\lambda+1}} - \frac{1}{(2z-1)^{\lambda+1}}\right) \ ,$$

and its first moment is:

$$\begin{aligned}E\left(Z_i | Z_i \geq Z_n\right) &= \int_1^{+\infty} z f\left(z\right) dz \\ &= 2\lambda \left[\int_1^{+\infty} \frac{1}{z^{\lambda}} dz - \int_1^{+\infty} \frac{z}{(2z-1)^{\lambda+1}} dz\right] \ .\end{aligned}$$

For the second integral inside the square brackets, with the substitution $2z - 1 = x$, we obtain

$$\int_1^{+\infty} \frac{z}{(2z-1)^{\lambda+1}} dz = \frac{1}{4}\int_1^{+\infty} \frac{1}{x^{\lambda}} dx + \frac{1}{4}\int_1^{+\infty} \frac{1}{x^{\lambda+1}} dx \ .$$

Hence:

$$\begin{aligned}E\left(Z_i | Z_i \geq Z_n\right) &= 2\lambda \left(\frac{3}{4}\int_1^{+\infty} \frac{1}{z^{\lambda}} dz - \frac{1}{4}\int_1^{+\infty} \frac{1}{z^{\lambda+1}} dz\right) \\ &= \frac{2\lambda + 1}{2\left(\lambda - 1\right)} \\ &= E\left(Z_i\right) + \frac{1}{2\left(\lambda - 1\right)} \ ,\end{aligned}$$

and the TFP gain immediately obtains by dividing by $E\left(Z_i\right)$.

D.3 Normal

If (Z_1, Z_2) has a bivariate normal distribution with the mean and variance of Z_i equal to μ_i and σ_i^2 $(i = 1, 2)$, with $\sigma_1^2 = \sigma_2^2 = s^2$, and with the correlation between Z_1 and Z_2 given by ρ, then it is easy to verify a useful property: the variables $U = Z_1 + Z_2$ and $V = Z_1 - Z_2$ are normally distributed and independent from each other.[37] Note also that the standard deviation of V is $\sigma_v = s\sqrt{2(1-\rho)}$.

Hence, we can write $Z_1 = (U+V)/2$, while $Z_1 \geq Z_2$ is equivalent to $V \geq 0$. Therefore:

$$\begin{aligned} E(Z_1|Z_1 \geq Z_2) &= E\left(\frac{U+V}{2}|V \geq 0\right) = \frac{1}{2}E(U) + \frac{1}{2}E(V|V \geq 0) = \\ &= \frac{\mu_1+\mu_2}{2} + \frac{1}{2}\left[\mu_1 - \mu_2 + \frac{\sigma_v \cdot g\left(\frac{\mu_2-\mu_1}{\sigma_v}\right)}{1 - G\left(\frac{\mu_2-\mu_1}{\sigma_v}\right)}\right], \end{aligned}$$

where the last step follows from the properties of normal and truncated normal random variables (with g and G respectively equal to the p.d.f. and c.d.f. of a standard normal random variable). After simplifying and dividing by μ_1, we immediately obtain:

$$\frac{E(Z_1|Z_1 \geq Z_2)}{E(Z_1)} = 1 + \frac{\sigma_v}{2\mu_1}\frac{g\left(\frac{\mu_2-\mu_1}{\sigma_v}\right)}{1 - G\left(\frac{\mu_2-\mu_1}{\sigma_v}\right)}.$$

The stronger result that $Z_i|Z_i \geq Z_n$ first-order stochastically dominates Z_i follows from the properties of normal distributions and the fact that the former variable has a higher mean and a lower variance.

D.4 Lognormal

If $Y_1 = \exp(Z_1)$, $Y_2 = \exp(Z_2)$, and $(Z_1, Z_2) \sim Norm\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$ — where, for simplicity, we take $\sigma_1^2 = \sigma_2^2 = s^2$ — then, by definition, (Y_1, Y_2) is a bivariate lognormal distribution.

Using the same variable $U = Z_1 + Z_2$ and $V = Z_1 - Z_2$ built above, we have:

$$\begin{aligned} E(Y_1|Y_1 \geq Y_2) &= E\left(\exp\left(\frac{U+V}{2}\right)|V \geq 0\right) \\ &= E\left(\exp\left(\frac{U}{2}\right)\right)E\left(\exp\left(\frac{V}{2}\right)|V \geq 0\right) \\ &= E\left(\exp\left(\frac{U}{2}\right)\right)E\left(\exp\left(\frac{V}{2}\right)|\exp\left(\frac{V}{2}\right) \geq 1\right) \\ &\geq E\left(\exp\left(\frac{U}{2}\right)\right)E\left(\exp\left(\frac{V}{2}\right)\right) \\ &= E(Y_1). \end{aligned}$$

[37] It is easy to check that $Cov(U, V) = 0$. Since (U, V) is bivariate normally distributed, then uncorrelation implies independence.

To compute the conditional mean, consider that since the variable $Z_1|Z_1 \geq Z_2$ is normal, then the variable $\exp(Z_1)|Z_1 \geq Z_2$ is lognormal. Therefore:

$$\begin{aligned} E\left(Y_1|Y_1 \geq Y_2\right) &= E\left[\exp(Z_1)|Z_1 \geq Z_2\right] \\ &= \exp\left(\bar{\mu}_1 + \frac{1}{2}\bar{\sigma}_1^2\right) \ , \end{aligned}$$

where $\bar{\mu}_1 = E\left(Z_1|Z_1 \geq Z_2\right)$ and $\bar{\sigma}_1^2 = Var\left(Z_1|Z_1 \geq Z_2\right)$. In particular:

$$\bar{\mu}_1 = \mu_1 + \frac{\sigma_v}{2}\lambda\left(\alpha\right) \ ,$$

where $\sigma_v = s\sqrt{2\left(1-\rho\right)}$ and $\alpha = \left(\mu_2 - \mu_1\right)/\sigma_v$ and λ is the hazard function of a standard normal distribution. For $\bar{\sigma}_1^2$, using the variables U and V once again we have:

$$\begin{aligned} \bar{\sigma}_1^2 &= Var\left(\frac{U+V}{2}|V \geq 0\right) \\ &= Var\left(\frac{U}{2}\right) + Var\left(\frac{V}{2}|V \geq 0\right) \\ &= \frac{s^2}{2}\left(1+\rho\right) + \frac{s^2}{2}\left(1-\rho\right)\left[1-\delta\left(\alpha\right)\right] \\ &= s^2\left[1 - \frac{1-\rho}{2}\delta\left(\alpha\right)\right] \ , \end{aligned}$$

where:

$$\delta\left(\alpha\right) = \lambda\left(\alpha\right)\left[\lambda\left(\alpha\right) - \alpha\right] \ .$$

As above, the stronger result that $Z_i|Z_i \geq Z_n$ first-order stochastically dominates Z_i follows from the properties of lognormal distributions and the fact that the former variable has a higher mean and a lower variance.

E Weighted TFP

E.1 Closed and open economies

We start this section by considering the weighted TFP of the closed economy, proving Proposition 3. Recall, from equation (13), that $\omega_i\left(z_i\left(j\right)\right) = p_i^{\sigma-1}/p_i^{\sigma-1}\left(j\right)$. Prices are equal to marginal costs, then $p_i^{\sigma-1}\left(j\right) = w_i^{\sigma-1}/z_i^{\sigma-1}\left(j\right)$. In addition, from the definition of p_i we find:

$$p_i^{\sigma-1} = \frac{w_i^{\sigma-1}}{E\left(Z_i^{\sigma-1}\right)} \ . \qquad (21)$$

Therefore: $\omega_i\left(z_i\left(j\right)\right) = z_i^{\sigma-1}\left(j\right)/E\left(Z_i^{\sigma-1}\right)$. By substituting it into equation (12), Proposition 3 immediately obtains.

We now turn to the open economy and consider an example with two countries, n and i, and defer a brief discussion about the extension to N countries to the next section. For the sake of simplicity, we neglect intermediate ($\beta = 1$) and non-tradeable goods (then $Y_i = w_i L_i$). With finite trade barriers, we need to consider separately non-exported and exported goods.

If the good j is not exported, then the demand $c_i(j)$ is the same as in equation (11) (where, however, in the open economy the price index p_i depends also on foreign technologies, because some goods are imported). Hence, the weight of good j is the same as in equation (13). In particular, equation (21) still holds, with $E(Z_i^{\sigma-1})$ replaced by $E(Z_{i,o}^{\sigma-1})$. Note that, even though some goods are imported, the price index of country i depends solely on the productivities of country i's industries. This result follows from a property of the EK model, that the prices of the goods that country i actually buys from any other country (including i itself) is the same as the overall price distribution of country i (i.e. when a country starts selling in country i, it does that up to the point at which the price distribution of the goods that it sells in i is the same as the overall price distribution of i).

If the good j is exported, then it is demanded by both countries i and n. The weight of good j becomes: $\omega_i(z_i(j)) = p_i(j)\, c_w(j)\, Y_i^{-1} = L_i(j)\, L_i^{-1}$, where $c_w(j) = c_i(j) + c_n(j)$ is the total (world) demand for good j. Recalling that the price of good j produced by country i and sold in the destination market n is $p_{ni}(z) = p_i(j)\, d_{ni}$, we can break the weight $\omega_i(z_i(j))$ into two different addenda: $p_i(j)\, c_i(j)\, Y_i^{-1}$ and $p_{ni}(j)\, c_n(j)\, (d_{ni} Y_i)^{-1}$. The former addendum is given by equation (21), with $E(Z_i^{\sigma-1})$ replaced by $E(Z_{i,o}^{\sigma-1})$. The latter, that depends on demand in country n, is: $p_n^{\sigma-1} p_{ni}^{1-\sigma}(j)\, Y_n (d_{ni} Y_i)^{-1}$. In particular, from equation (2) it follows that: $p_n = p_i (\Phi_i/\Phi_n)^{1/\theta}$.

Hence, the weight of any good j produced by country i becomes:

$$\omega_i(j) = \begin{cases} \dfrac{z_i^{\sigma-1}(j)}{E\left(Z_{i,o}^{\sigma-1}\right)} & \text{if } j \text{ is not exported} \\ \dfrac{z_i^{\sigma-1}(j)}{E\left(Z_{i,o}^{\sigma-1}\right)}(1+\kappa) & \text{if } j \text{ is exported} \end{cases} ,$$

where:

$$k = \frac{Y_n}{Y_i}\left(\frac{\Phi_i}{\Phi_n}\right)^{(\sigma-1)/\theta} d_{ni}^{-\sigma} .$$

Thus, we are ready to compute the weighted TFP of the open economy:

$$\mathrm{TFP}_{i,o}^{w} = \int_0^{+\infty} z_i(j)\, \omega_i(j)\, dF_{i,o}(j) .$$

Anticipating the result developed in the next section that:

$$F_{i,o}(z) = (1-\lambda)\, F_{i,d}(z) + \lambda F_{i,e}(z) , \tag{22}$$

where $F_{i,d}$ ($F_{i,e}$) is the c.d.f. of the productivity distribution of non-exporters (exporters) and:

$$\lambda = \frac{\Phi_i}{\Phi_n} d_{ni}^{-\theta} , \tag{23}$$

we can compute:

$$\begin{aligned}
\mathrm{TFP}_{i,o}^{w} &= (1-\lambda)\int_{0}^{+\infty} z_i(j)\,\omega_i(j)\,dF_{i,d}(j)+\lambda\int_{0}^{+\infty} z_i(j)\,\omega_i(j)\,F_{i,e}(j)\\
&= \frac{1}{E\left(Z_{i,o}^{\sigma-1}\right)}\left[(1-\lambda)\int_{0}^{+\infty} z_i^{\sigma}(j)\,dF_{i,d}(j)+(\lambda+\lambda\kappa)\int_{0}^{+\infty} z_i^{\sigma}(j)\,dF_{i,e}(j)\right]\\
&= \frac{1}{E\left(Z_{i,o}^{\sigma-1}\right)}\left[\int_{0}^{+\infty} z_i^{\sigma}(j)\,dF_{i,o}(j)+\lambda\kappa\int_{0}^{+\infty} z_i^{\sigma}(j)\,dF_{i,e}(j)\right]\\
&= \frac{E\left(Z_{i,o}^{\sigma}\right)}{E\left(Z_{i,o}^{\sigma-1}\right)}+\lambda\kappa\frac{E\left(Z_{i,e}^{\sigma}\right)}{E\left(Z_{i,o}^{\sigma-1}\right)}\;.
\end{aligned}$$

This results proves equation (15), with $b=\lambda\kappa$. Using equation (2) (as well as equation (12) in EK), notice that the "extra weight" given to the exporters is:

$$b=\lambda\kappa=\frac{Y_n/p_n}{Y_i/p_i}\left(\frac{X_{ni}/X_n}{X_{ii}/X_i}\right)^{(\sigma+\theta)/\theta}\;.$$

As the country approaches autarky ($d_{in},d_{ni}\to+\infty$), then $\lambda\to 0$, $Z_{i,d}\to Z_{i,o}$ and, in turn, $Z_{i,o}\to Z_i$; in other words, $\mathrm{TFP}_{i,o}^{w}\to\mathrm{TFP}_{i}^{w}$. As the country approaches free trade ($d_{in},d_{ni}\to 1$), then $\lambda\to 1$, $Z_{i,e}\to Z_{i,o}$ because the set of non-exported goods becomes empty, and $\mathrm{TFP}_{i,o}^{w}\to E\left(Z_{i,o}^{\sigma}\right)/E\left(Z_{i,o}^{\sigma-1}\right)$.

E.2 Productivity distribution of exporters and non-exporters

In this section, we compute the distribution of productivities for exporters and non-exporters and prove equation (22). As above, we consider an example with two countries, neglecting intermediate and non-tradeable goods. Recall that country i produces good j if and only if $z_i(j)>z_n(j)\,w_i\,[w_n d_{in}]^{-1}$. Among these surviving industries, good j is: (i) exported $\Leftrightarrow$ $z_i(j)>z_n(j)\,w_i d_{ni} w_n^{-1}$; ($ii$) not exported $\Leftrightarrow$ $z_n(j)\,w_i\,[w_n d_{in}]^{-1}<z_i(j)<z_n(j)\,w_i d_{ni} w_n^{-1}$.

Assuming that $Z_i\sim \textit{Fréchet}\,(T_i,\theta)$, Proposition 1 provides the distribution of productivities for the industries that survive international competition in the general case. In our simplified example, this is a Fréchet distribution with state of technology $T_i+T_n w_i^{\theta}(w_n d_{in})^{-\theta}$ and precision θ.

Similarly, we can compute the c.d.f. $F_{i,e}$ of the distribution of the exporters as:

$$F_{i,e}(z)=\Pr\left(Z_i<z|Z_i>Z_n w_i d_{ni} w_n^{-1}\right)=\frac{\Pr\left(Z_i<z,Z_i>Z_n w_i d_{ni} w_n^{-1}\right)}{\Pr\left(Z_i>Z_n w_i d_{ni} w_n^{-1}\right)}\;.$$

Solving the corresponding integrals, we find that $Z_{i,e}$ is Fréchet distributed, with state of technology $T_i+T_n(w_i d_{ni})^{\theta} w_n^{-\theta}$ and precision θ. Note, in particular, that $E(Z_{i,e})>E(Z_{i,o})$ and that, as the country approaches free trade , then $Z_{i,e}\to Z_{i,o}$.

Denoting by $Z_{i,d}$ the distribution of the producers who sell their goods only domestically (non-exporters) and by $F_{i,d}$ its c.d.f., we have:

$$\begin{aligned} F_{i,d}(z) &= \Pr\left(Z_i < z | Z_n w_i (w_n d_{in})^{-1} < Z_i < Z_n w_i d_{ni} w_n^{-1}\right) \\ &= \frac{\Pr\left(Z_i < z, Z_i > Z_n w_i (w_n d_{in})^{-1}\right) - \Pr\left(Z_i < z, Z_i > Z_n w_i d_{ni} w_n^{-1}\right)}{\Pr\left(Z_n w_i (w_n d_{in})^{-1} < Z_i < Z_n w_i d_{ni} w_n^{-1}\right)} \ . \end{aligned}$$

In few steps we can obtain:

$$F_{i,o}(z) = \frac{\Pr\left(Z_n w_i (w_n d_{in})^{-1} < Z_i < Z_n w_i d_{ni} w_n^{-1}\right) \cdot F_{i,d}(z) + \Pr\left(Z_i > Z_n w_i d_{ni} w_n^{-1}\right) \cdot F_{i,e}(z)}{\Pr\left(Z_i > Z_n w_i (w_n d_{in})^{-1}\right)}$$

from which equation (22) follows. In other words, $F_{i,o}(z)$ is a mixture between the distribution of non-exporters and the distribution of exporters, with weights respectively equal to $1 - \lambda$ and λ, where λ is given by equation (23)

In a more general setting (with $N > 2$), we can find a closed-form expression for the distribution of $Z_{i,e}$ (that appears in equation (15)), but the weighting scheme $\omega_i(j)$, which is needed to compute $\mathrm{TFP}^w_{i,o}$, becomes increasingly cumbersome. Let us first compute $Z_{i,e}$. Denote by $Z_{i,e,n}$ the productivity distribution for the goods that country i exports to country n. To generalize condition (i) above, note that country i exports good j in country $n \Leftrightarrow z_i(j) > w_i d_{ni} \max_{k \neq i} [z_k(j) / (w_k d_{nk})]$. It is easy to check that $Z_{i,e,n}$ is Fréchet distributed, with state $T_i + \sum_{k \neq i} T_k (w_i d_{ni} / w_k d_{nk})^\theta$ and precision θ. By the same token, note that good j is exported if and only if the following condition holds:

$$z_i(j) > \min_{n \neq i} \left\{ w_i d_{ni} \max_{k \neq i} \left[\frac{z_k(j)}{w_k d_{nk}} \right] \right\} \ .$$

In other words, $Z_{i,e} = \min_{n \neq i}(Z_{i,e,n})$; therefore:

$$F_{i,e}(z) = 1 - \prod_{n \neq i} [1 - F_{i,e,n}(z)] \ ,$$

where $F_{i,e,n}$ is the c.d.f. of $Z_{i,e,n}$. As we expand the product, we obtain:

$$F_{i,e}(z) = \sum_{n \neq i} F_{i,e,n}(z) - \sum_{n \neq i, k < n} F_{i,e,n}(z) F_{i,e,k}(z) + ... + (-1)^N \prod_{n \neq i} F_{i,e,n}(z) \ .$$

Clearly, $Z_{i,e}$ is composed by the productivity distributions for the goods that country i exports in country n ($n \neq i$), the productivity distributions for the goods that i exports in countries n and k ($n \neq i$ and $k < n$), etc.. Overall, this a mixture of $2^{N-1} - 1$ Fréchet distributions; with $N = 41$, as in our empirical analysis, it would be a mixture of about 10^{12} distributions.

What about the weights needed to compute $\mathrm{TFP}^w_{i,o}$? In general, each good j has a different weight, depending on whether and where it is exported. For any good j, we can write: $\omega_i(j) \propto z_i^{\sigma-1}(j)(1 + \kappa(j))$, where the "extra weight" $\kappa(j)$ depends on the destination market of each good j (as just noticed, with N countries, we have $2^{N-1}-1$ different possible destinations for exports). When we substitute $\omega_i(j)$ into $\text{TFP}^w_{i,o}$, we still get two terms (like in equation (15)). The first is $E(Z^{\sigma}_{i,o})/E(Z^{\sigma-1}_{i,o})$, and reflects the "pure" selection effect; the second, which is more complicated because weights have to be attributed to each possible destination market, reflects the market-share reallocation effect.

F Data sources

Our main data source is OECD STAN 2008, which provides data on production and trade for the manufacturing sector as a whole. For the countries not covered by STAN, we use data from UNIDO IDSB 2008, from which we recover production and trade of the manufacturing sector by aggregating data available for all manufacturing industries at the 4-digit Level of ISIC classification (Revision 3 for the period 1990-2005, Revision 2 for 1985-1989). For the major industrial countries, which are included in both databases, results from using one or the other source are essentially the same.

For Belgium, which is covered by STAN, we use IDSB data because STAN provides the inconsistent result that Belgian manufacturing exports are larger than its manufacturing production, an artifact of its role as entrepôt country. Using disaggregated data from IDSB, then, we can neglect manufacturing industries in which production is either zero or missing, a simple method that is sufficient to return exports that are smaller than production. We cannot exclude, however, that entrepôt trade inflates export and import flows also for other sectors where Belgian production is reported, as well as for other countries.

Finally, notice that in the theoretical model, the budget constraint implies that exports and imports are equal.[38] Therefore, in equation (6) one could use exports and imports interchangeably. For this reason, we have computed the values reported in Table 1 by replacing EXP_i and IMP_i with the mean $(EXP_i + IMP_i)/2$.

[38] For an extension of the model to the case of "unbalanced trade", see Dekle, Eaton, and Kortum (2007).

References

[1] Alcalá F., Ciccone A. (2004), "Trade and Productivity," *Quarterly Journal of Economics*, Vol. 119, pp. 612-645.

[2] Alvarez F., Lucas R.E.Jr. (2007), "General Equilibrium Analysis of the Eaton-Kortum Model of International Trade," *Journal of Monetary Economics*, Vol. 54, pp. 1726-1768.

[3] Anderson J.E., van Wincoop E. (2004), "Trade costs," *Journal of Economic Literature*, Vol. 42, pp. 691–751.

[4] Armington P.S. (1969), "A Theory of demand for Products Distinguished by Place of Production," *IMF Staff Papers*, No. 16.

[5] Balakrishnan N., Nevzorov V.B. (2003), *A Primer on Statistical Distributions*, Wiley, Hoboken, New Jersey.

[6] Bernard A.B., Eaton J., Jensen J.B., Kortum S. (2003), "Plants and Productivity in International Trade," *American Economic Review*, Vol. 93, pp. 1268-1290.

[7] Bernard A.B., Jensen J.B. (1999), "Exceptional Exporter Performance: Cause, Effect, or Both?," *Journal of International Economics*, Vol. 47, pp. 1-25.

[8] Bernard A.B., Jensen J.B., Redding S.J., Schott P.K. (2007) "Firms in International Trade," *Journal of Economic Perspectives*, Vol. 21, pp. 105-130.

[9] Chaney T. (2008), "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, Vol. 98, pp. 1707-1721.

[10] Chari V.V., Restuccia D., Urrutia C. (2005), "On-the-Job Training, Limited Commitment, and Firing Costs," *unpublished*, University of Minnesota.

[11] Conway P., Nicoletti G. (2006), "Product Market Regulation in the Non-Manufacturing Sectors of OECD Countries: Measurement and Highlights," *OECD Economics Department Working Papers*, No. 530.

[12] Demidova S., Rodríguez-Clare A. (2009), "Trade Policy under Firm-Level Heterogeneity in a Small Economy," *Journal of International Economics*, forthcoming.

[13] Dekle R., Eaton J., Kortum S. (2007), "Unbalanced Trade," *American Economic Review*, Vol. 97, pp. 351-355.

[14] Dollar D., Kraay A. (2003), "Institutions, trade, and growth," *Journal of Monetary Economics*, Vol. 50, pages 133-162.

[15] Dornbusch R., Fischer S., Samuelson P. A. (1977), "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods," *American Economic Review*, Vol. 67, pp. 823-839.

[16] Eaton J., Kortum S. (2002), "Technology, Geography, and Trade," *Econometrica*, Vol. 70, pp. 1741-1779.

[17] Eaton J., Kortum S. (2009), *Technology in the Global Economy: A Framework for Quantitative Analysis*, Princeton University Press, forthcoming (title tentative).

[18] Frankel J.A., Romer D. (1999), "Does Trade Cause Growth?," *American Economic Review*, Vol. 89, pp. 379-399.

[19] Finicelli A., Pagano P., Sbracia M. (2009a), "Trade-revealed TFP," *unpublished*, Bank of Italy.

[20] Finicelli A., Pagano P., Sbracia M. (2009b), "The Eaton-Kortum Model: Empirical Issues and Extensions," *unpublished*, Bank of Italy.

[21] Gumbel E.J. (1960), "Bivariate Exponential Distributions," *Journal of the American Statistical Association*, Vol. 55, pp. 698-707.

[22] Hall R., Jones C. I. (1999), "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, Vol. 114, pp. 83-116.

[23] Jones C.I. (2005), "The Shape of Production Functions and the Direction of Technical Change," *Quarterly Journal of Economics*, Vol. 120, pp. 517-549.

[24] Kortum S. (1997), "Research, Patenting, and Technological Change," *Econometrica*, Vol. 65, pp. 1385-1419.

[25] Lagos R. (2006), "A Model of TFP," *Review of Economic Studies*, Vol. 73, pp. 983-1007.

[26] Mardia K.V. (1970), *Families of Bivariate Distributions*, Lubrecht & Cramer (Griffin's Statistical Monographs, No 27).

[27] Melitz M.J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, Vol. 71, pp. 1695-1725.

[28] Melitz M.J., Ottaviano G.I.P. (2008), "Market Size, Trade, and Productivity," *Review of Economic Studies*, Vol. 75, pp. 295-316.

[29] Pavcnik N. (2002), "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *Review of Economic Studies*, Vol. 69, pp. 245-76.

[30] Rodríguez-Clare A. (2007), "Trade, Diffusion and the Gains from Openness," *NBER Working Papers*, No. 13662, National Bureau of Economic Research.

[31] Tawn J.A. (1990), "Modelling Multivariate Extreme Value Distributions," *Biometrika*, Vol. 77, pp. 245-253.

[32] Waugh M.E. (2008), "International Trade and Income Differences", *unpublished*, University of Iowa.